FIW – Working Paper

FIW Working Paper N° 45
January 2010

Agglomeration Economies and Location Choices by Foreign Firms in Vietnam

Dinh Thi Thanh Binh*

Abstract

This paper studies the effects of agglomeration economies on the location choices by foreign firms in Vietnam. By using a large dataset that provides detailed information about individual firms, the study examines the location choices by 568 newly created foreign firms in 2005 in about 150 different 4-digit industries. This is one of the few studies of agglomeration effects on the location choices by foreign investments in transition economies in general and in Vietnam in particular. The estimates of the negative binomial regression model and the conditional logit model show that agglomeration benefits motivate foreign firms in the same industries and from the same countries of origin to locate near each other. However, the empirical results also indicate that there is competition among provinces in Vietnam in attracting foreign investors, and the locations of Vietnamese firms have no effects on the location decisions by foreign firms in the same industry.

JEL Codes: F 23
Keywords: Agglomeration; Location choice; Foreign direct investment

The author

*Ph.D. student of the Doctoral School of Economics and Management, University of Trento, Italy
Mailing address: via Rosmini, 70, Trento (38100), Italy
Email: dinhdieubinh@yahoo.com

Agglomeration Economies and Location Choices by Foreign Firms in Vietnam

Dinh Thi Thanh Binh*

Ph.D. student of the Doctoral School of
Economics and Management, University of Trento, Italy
Mailing address: via Rosmini, 70, Trento (38100), Italy
Email: dinhdieubinh@yahoo.com
The birthday of the author: 10/03/1979

Abstract: This paper studies the effects of agglomeration economies on the location choices by foreign firms in Vietnam. By using a large dataset that provides detailed information about individual firms, the study examines the location choices by 568 newly created foreign firms in 2005 in about 150 different 4-digit industries. This is one of the few studies of agglomeration effects on the location choices by foreign investments in transition economies in general and in Vietnam in particular. The estimates of the negative binomial regression model and the conditional logit model show that agglomeration benefits motivate foreign firms in the same industries and from the same countries of origin to locate near each other. However, the empirical results also indicate that there is competition among provinces in Vietnam in attracting foreign investors, and the locations of Vietnamese firms have no effects on the location decisions by foreign firms in the same industry.

Key words: Agglomeration; Location choice; Foreign direct investment

JEL classification: F23

---

* I would like to thank my advisor, Marco Zamarian. I also would like to thank Klaus Meyer, Christopher Gilbert, and Marco Corsino for valuable comments and feedbacks at different stages of the paper. In addition, I would like to thank the Trento Chamber of Commerce and CIFREM for financial support. Any errors and omissions remain my own.

1. Introduction

According to traditional trade theory, location choice by a foreign firm depends on factor endowments of host countries such as natural resources, labor capital and infrastructures. The "factor endowment" theory, which was developed from Ricardo's theory of comparative advantages by Heckscher and Ohlin (Krugman and Obstfeld, 1997), claims that firms have tendencies to locate in places where the required factors of their production are relatively abundant. However, recent theories of economic geography suggest that firms in the same industries may be drawn to a particular location in order to benefit from positive externalities or agglomeration effects.

The theory of agglomeration economies was introduced by Marshall (1920) in which he provided three reasons for the clustering of firms in the same industries: it provides a pooled market for workers with specialized skills, facilitates the development of specialized inputs and services, and enables firms to benefit from technological spillovers. Subsequent research by Krugman (1991) and Saxenian (1994) construct formal models to analyze and extend the concepts.

To date, there have been few empirical studies on agglomeration effects, especially in transition economies. Head, Ries and Swenson (1995) examine location choices by Japanese firms in manufacturing industries in the United States, showing that Japanese firms prefer to locate near both US and Japanese firms in the same manufacturing industries. Guimaraes *et al.* (2000) and Crozet, Mayer and Mucchielli (2004) also indicate similar behavior by foreign firms in France and Portugal, respectively. However, there are also studies that do not support the existence of agglomeration effects. Shaver and Flyer (2000) examine foreign manufacturing firms in the United States and find that large firms are not likely to locate near other firms because the benefits they contribute to agglomeration economies are less than what they receive from agglomeration effects. Empirically, Baum and Mezias (1992) and Baun and Haveman (1997) also support this conclusion. For transition economies, there are fewer studies of agglomeration effects on location choices by foreign investors. Most important are the works of Boudier-Bensebaa (2005) on Hungary, Meyer and Nguyen (2005)[1] on Vietnam, and Head and Ries (1996) and Cheng and Kwan (2000) on China. However, due to the lack of detailed firm-level information, these studies can use only aggregate numbers of firms or foreign investment projects at provincial levels to estimate agglomeration effects.

[1] Meyer and Nguyen (2005) did not concentrate on agglomeration. Yet, the authors have a small data analysis and discussion about the effects of economic agglomeration on the location choices by foreign investors in Vietnam.

This study includes investments of 568 newly created foreign firms in 2005 in about 150 different 4-digit industries. We also controls for the effects of province-specific factor endowments by using provincial characteristics in the model and for the effect of industry-specific endowments by using the geographical patterns of 88420 Vietnamese firms in the same industries during 2004. The study shows that the deviation of foreign firms from these patterns indicates agglomeration effects. Different from many other studies, "country of origin" is used as a new dimension in the measurement of agglomeration effects.

We apply the negative binomial regression model and the conditional logit model to estimate the effects of agglomeration economies on location choices by newly created foreign firms in Vietnam in 2005. By using a large dataset and detailed information about individual firms, it is possible to measure the effects of the country of origin and the industry of a firm on its location choice. The study shows that foreign investors are not only likely to locate near other foreign firms but also prefer to locate near foreign firms in the same industries and from the same countries of origin. Similar to Head *et al.* (1995), it is argued that this pattern of location choice supports an agglomeration-externality theory rather than a theory based on the differences of endowment factors. Further, the empirical results reveal that there is competition among provinces in attracting foreign investors, and the locations of Vietnamese firms have no effect on the location decisions by foreign investors in the same industries.

This research contributes to the existing literature on agglomeration economies, location and foreign direct investment. To the best of our knowledge, this is the first study of agglomeration effects on the location choices by foreign investors in Vietnam using detailed information about individual firms. The empirical results are particularly important for Vietnam's provincial authorities in designing policies aimed at attracting foreign investments.

The structure of this paper is organized as follows. Section 2 provides an overview of regional economies and the stylized facts of the FDI patterns by provinces in Vietnam. Section 3 reviews theories on localization. Section 4 describes the dataset. Section 5 presents methodology and empirical results. The final section is devoted to conclusions.

2. An overview of regional economies and the stylized facts of the FDI pattern in Vietnam

Regional economies

Vietnam is divided into fifty-nine provinces and five centrally-controlled municipalities in eight regions based on geographical and socio-economic conditions. The eight regions are Red River Delta, Northeast, Northwest, North Central Coast, South Central Coast, Central Highlands, Southeast, and Mekong River Delta (see Figure 2). The Red River Delta, the Southeast, and the Mekong River Delta have much smaller areas compared with the others, but they are the most densely populated areas, accounting for 58.7% of the country's population in 2005. By contrast, the Northwest and the Central Highlands are the least populated regions with less than 9% of the country's population in 2005 (see Table 1).

Table 1: General indicators of the regions in Vietnam

Region	Population share 2005 (%)	Agricultural share 2005 (%)	Industrial share 2004 (%)	Service share 2005 (%)	Income per capita 2004 (thousand VND)
Red River Delta	21.7	17.6	19.2	19.9	5858.4
Northeast	11.3	8.1	4.5	6.2	4558.8
Northwest	3.1	2.2	0.2	1.1	3188.4
North Central Coast	12.8	8.5	2.4	6.1	3805.2
South Central Coast	8.5	5.2	4.0	7.8	4978.8
Central Highlands	5.7	11.8	0.6	3.4	4682.4
Southeast	16.2	11.7	57.1	36.3	9996.0
Mekong River Delta	20.8	35.0	8.0	19.3	5653.2

Source: The Statistical Yearbook of Vietnam in 2005.

Note: The agricultural output value is at constant 1994 prices, the other indicators are at current prices.

The Red River Delta including Hanoi, the capital and the Southeast including Ho Chi Minh City, the largest city of Vietnam are also the most developed regions in Vietnam. These regions are the major industrial centers of the country, producing 19.2% and 57.1% respectively of the country's industrial output in 2004. The Northwest and the Central Highlands, on the other hand, are the least industrialized regions with industrial output less than 1% of the nation's total in 2004 (The Statistical Yearbook of Vietnam in 2005).

Regarding agricultural production, the Mekong River Delta and the Red River Delta are the two major rice-producing areas in Vietnam, accounting for 52.6% of the country's agricultural output in 2005. The Southeast, the Mekong River Delta, and the Red River Delta are also the most important centers for services in Vietnam,

and they have the three largest cities of Ho Chi Minh City, Can Tho, and Hanoi, respectively. Those regions accounted for 75.5% of the country's total service output in 2005 (see Table 1).

As a result of being the biggest centers in agriculture, industry, and services, the living standards of people in the South East, the Red River Delta, and the Mekong River are the highest in Vietnam.

The FDI pattern

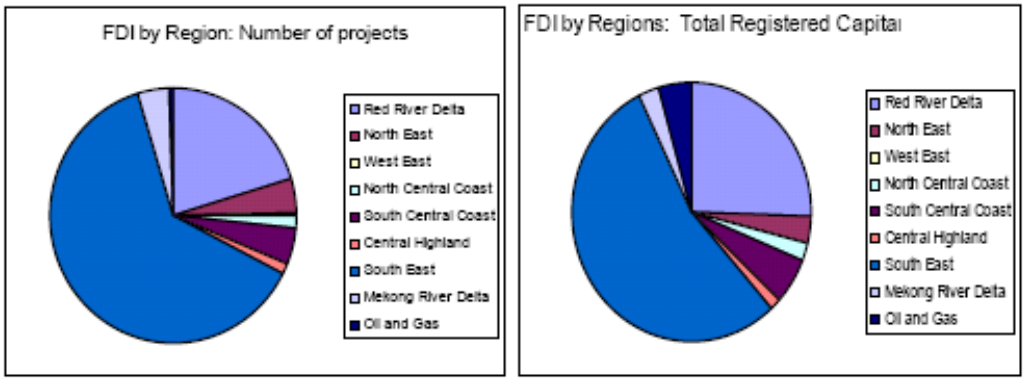
Since the Law of Foreign Investment was passed in 1987, the flows of FDI into Vietnam have been considerable and have also increased over time. However, the increasing trend has not been smooth. After a big jump during the period 1988-1996, Vietnam experienced a sharp decline in FDI flows at the final years of 1990s due to strong influence of the Asian financial crisis in 1997. However, the FDI inflows started to pick up again as countries in the region recovered from the crisis and the United States-Vietnam Bilateral Trade Agreement was signed in 2001. Especially, the situation has changed much since Vietnam became a formal member of the WTO in the beginning of 2007. According to the Ministry of Planning and Investment of Vietnam (MPI), in 2007 FDI inflows into Vietnam achieved the highest record with $21.3 billion of registered capital after twenty years of issuing the first Law on Foreign Direct Investment.

The statistic data of the MPI show an uneven distribution of FDI in both industrial sectors and regions during the period 1988-2007 by the number of investment projects and the amount of registered capital. In terms of industrial sector, nearly 70% of projects and registered capital were running to manufacture, around 20% to service and the rest to agriculture. Within the manufacture, while during the early part of 1990s, the majority of FDI were in oil and mining sector, but recently light and heavy industries dominate the field. In addition, the share of FDI in agricultures now is increasing compared with that in the 1990s. In service sector, the hotel and tourism activities account for the largest proportion. A different point is that in the early history of the FDI in Vietnam, in the service sector, there was no investment in construction of industrial zones, offices and apartments, but now these fields start attracting significant part of FDI inflows.

In terms of nationalities of investors, the data of the MPI reveals that during 1988-2007, there were eighty one countries and territories investing in Vietnam. The inward FDI in Vietnam is dominated by regional investors, accounting for nearly 80% of the total number of investment projects, registered capital and implemented capital. The top five investors were South Korea, Taiwan, Japan,

China, and Singapore. Although the United States is a late comer, its investment in Vietnam has increased since the Bilateral Trade Agreement between the two countries was signed, and now it is in the eighth position of investment ranking. The investments from European countries were still small, accounting for about 10% of the numbers of projects, 15% of the registered capital and 20% of the implemented capital.

Figure 1: The regional distributions of FDI in Vietnam during 1988-2007



Source: The MPI

Regarding regional distribution, during the period 1988-2007, all sixty four provinces in Vietnam had received FDI, but most of them flew to the Southeast and the Red River Delta regions. Figure 1 shows that more than 60% of projects and 52% of registered capital ran to the Southeast region of which most of them flew to Ho Chi Minh City and its two neighboring provinces, Dong Nai and Binh Duong, and nearly 25% of investment projects and registered capital went to the Red River Delta of which Hanoi, the capital city, accounted for the largest proportion. By contrast, the Northwest and the North Central Coast attracted less than 1% of the FDI inflows.

It seems that there is a relationship between the regional economic conditions and the FDI inflows. Most of the empirical studies on the distribution of the FDI in Vietnam show that market potential, labor force, infrastructure, agglomeration effects and institutional performance by local authorities are important determinants of FDI inflows into regions within Vietnam (Meyer and Nguyen, 2005; Nguyen Phuong Hoa, 2002; Pham Hoang Mai, 2002; Le Viet Anh, 2004; Nguyen Ngoc Anh and Nguyen Thang, 2007; Nguyen Phi Lan, 2006). However, besides conventional determinants of FDI location, recent theories of economic geography suggest that benefits arising from agglomeration economies drive foreign firms to locate in a particular place, therefore affecting FDI inflows. In the next section, we first review

the theories that explain agglomeration economies and then we advance three hypotheses of this study.

3. Theories of localization

Industry localization is defined as "the geographic concentration of particular industries" (Head *et al.*, 1995). One of the mechanisms motivating this concentration is the existence of agglomeration economies, which are positive externalities that stem from the geographic clustering of industries. In this context, firms contribute to the externalities and also benefit from the externalities (Shaver and Flyer, 2000).

The issue on industry localization attracted the attention of economists in the late nineteenth century. The work of Marshall (1920) is considered as an early and influential economic analysis on this phenomenon. Marshall identifies three externalities that stem from industry localization: (i) localization enables firms to benefit from technological spillovers, (ii) localization provides a pooled market for workers with specialized skills that benefits both workers and firms, and (iii) localization creates a pool of specialized intermediate inputs for an industry in greater variety and at lower cost. These positive externalities have the potential to enhance the performance by firms that agglomerate.

According to Krugman (1991), the concept of technological spillovers is quite vague and general but it is the most frequently mentioned as a source of agglomeration effects. Useful information can flow between near firms, designers, engineers, and managers. For foreign companies, the spillovers of information can be the flows of experience-based knowledge about how to operate efficiently in the host countries (Head *et al.*, 1995). Many authors use such clusters as California's Silicon Valley and Boston's Route 128 to show that technological externalities are the most obvious reason for firms to agglomerate (Krugman, 1991; Saxenian, 1994). However, by contrast with the labor pooling or intermediate goods supply that are in principle measurable, technological spillovers can be invisible and difficult to measure. It can therefore be difficult to state clearly that either technological spillovers or specialized labor play a more important role in creating high-technological clusters, for instance in Silicon Valley and the high-fashion cluster in Milan.

As anticipated by Marshall (1920), localized industry allows a pooled market for workers with specialized skills to benefit both workers and firms. David and Rosenbloom (1990) argue that an increased number of firms reduce the possibility that a worker will be unemployed for a long time. Finally, this also benefits firms by increasing the supply of specialized employees and reducing the risk of high-wage

requirements from labor. Popular examples of this phenomenon are microelectronic manufacture in Silicon Valley (Saxenian, 1994) and carpet manufacture in Dalton, Georgia (Krugman, 1991).

Krugman (1991) argues that the combination of scale economies and transportation costs will motivate the users and suppliers of intermediate inputs to cluster near each other. Such agglomerations reduce the total transportation costs and make large centers of production become more efficient and have more diverse suppliers than small ones. This will encourage firms in the same industries to concentrate in one location. Krugman points out that a historical accident makes a firm locate in a particular place, and then the cumulative location choices allow such an accident to influence the long-run geographical pattern of industry.

From these observations, it seems that firms benefit from geographical localization when agglomeration economies exist. So far, there have been two types of studies that support the existence of agglomeration benefits. The first is qualitative studies of agglomerations that identify the existence of industry clusters and document the existence of agglomeration externality mechanism (Krugman, 1991; Saxenian, 1994). The second is empirical studies that try to find whether a firm has benefits when locating near other firms in the same industry or from the same country of origin. For example, the empirical research of Head *et al.* (1995), Head and Ries (1996), Head, Ries and Swenson (1999), Crozet *et al.* (2004), Guimaraes *et al.* (2000), and Coughlin and Segev (2000) find that firms in the same industries and from the same countries of origin have tendencies to locate near each other. However, the empirical study of Shaver and Flyer (2000) shows that under the existence of agglomeration economies, many firms will perform better if they do not cluster. These authors argue that firms not only capture benefits from agglomeration economies but also contribute to agglomeration economies. Therefore, large firms with the greatest capacity in technologies, human capital, training programs, suppliers, and distributors will try to locate away from their competitors because the benefits they gain from locating near their competitors will be less than what the competitors gain from them.

The problems firms will experience when participating in an industrial cluster can be the spillover of technology, employee defection to competitors, and the sharing of distributors and suppliers with neighboring firms. Yoffie (1993) shows that semiconductor managers decide to locate far from their competitors due to their concern that their technology might spill over to the near firms. Baum and Mezias (1992) indicate that locating closer to other hotels in Manhattan increases the survival chance of a hotel, but this benefit of agglomeration diminishes when hotel districts become crowded, pushing up prices and exacerbating competition.

In this study, based on the FDI patterns in Vietnam, three hypotheses aimed at verifying the existence of agglomeration economies are tested. The empirical research on different countries – see the studies of Boudier-Bensabaa (2005) on Hungary, Meyer and Nguyen (2005) on Vietnam, Head and Ries (1996) and Cheng and Kwan (2000) on China, Crozet *et al.* (2004) on France, and Guimaraes *et al.* (2000) on Portugal – show that new foreign firms are likely to locate near other foreign investors. By doing that, they may use the experience and performance by earlier investors as indicators of the underlying business climate at the location. Hence, it is possible to expect an empirical relationship between the location choice by a new foreign firm and the prior number of foreign firms in a particular province.

Hypothesis 1: *the greater the number of foreign firms already established in a province, the more likely new foreign investors are to invest in that province.*

In the case of Vietnam, as presented in section 2, there is an uneven distribution of foreign investments. It is proposed that the provinces that already have a lot of foreign investment will be more attractive to new foreign investors due to agglomeration effects. Following the work of previous authors (Boudier-Bensabaa, 2005; Meyer and Nguyen, 2005; Cheng and Kwan, 2000), the stock number of foreign investors at provincial level in the previous year is used as a proxy for foreign-specific agglomeration.

When studying the behavior by Japanese firms in the United States, Head *et al.* (1995; 1999) find that new Japanese firms prefer to locate near both Japanese and US firms in the same industries. Moreover, Japanese firms are likely to locate near Japanese firms in the same manufacturer-led *keiretsu*[2]. Crozet *et al.* (2004) also find similar evidence about the industrial concentrations of foreign firms in France. It seems that the benefits from technological spillovers, specialized labor markets, and the availability of input suppliers to the industry motivate firms in the same industries to cluster. Based on the empirical results of previous studies, the following hypothesis is advanced.

Hypothesis 2: *the greater the number of domestic firms and foreign firms in a specific industry already located in a province, the more likely new foreign investors in that industry are to locate in that province.*

[2] Keiretsu can be considered as industrial or vertical groups, i.e. those headed by large manufacturing companies whose members consist largely of component suppliers.

In order to test this hypothesis, it is proposed that new foreign firms have a tendency to locate in the provinces where many Vietnamese firms and other foreign firms in the same industries already existed. The lagged stock number of Vietnamese firms and foreign firms in the same industries by province are used as proxies for industry-specific agglomeration.

Besides finding that foreign firms are likely to locate near firms in the same industries, Head *et al.* (1995; 1999) and Crozet *et al.* (2004) also show that foreign firms prefer to locate near firms from the same countries of origin. Head *et al.* (1999) argue that agglomeration effects between Japanese firms may arise due to their different characteristics from the firms of other countries. For example, the preference for higher skilled workers because of a stronger desire for quality control or greater use of complex machinery might motivate a new Japanese firm to locate near earlier arrivals to be able to hire away employees trained in Japanese methods. Thus, it is possible to expect an empirical relationship between location choice by a new foreign firm and the prior number of foreign firms from the same countries of origin in a particular province.

Hypothesis 3: *the greater the number of foreign firms from a specific country already located in a province, the more likely new foreign investors from that country are to locate in that province.*

Based on the location patterns of foreign investors in Vietnam, it is proposed that foreign investors from the same countries of origin are likely to concentrate in a particular region. Following the work of Crozet *et al.* (2004), the lagged stock number of foreign firms from the same countries of origin by province is used as a proxy for country-specific agglomeration.

4. Data

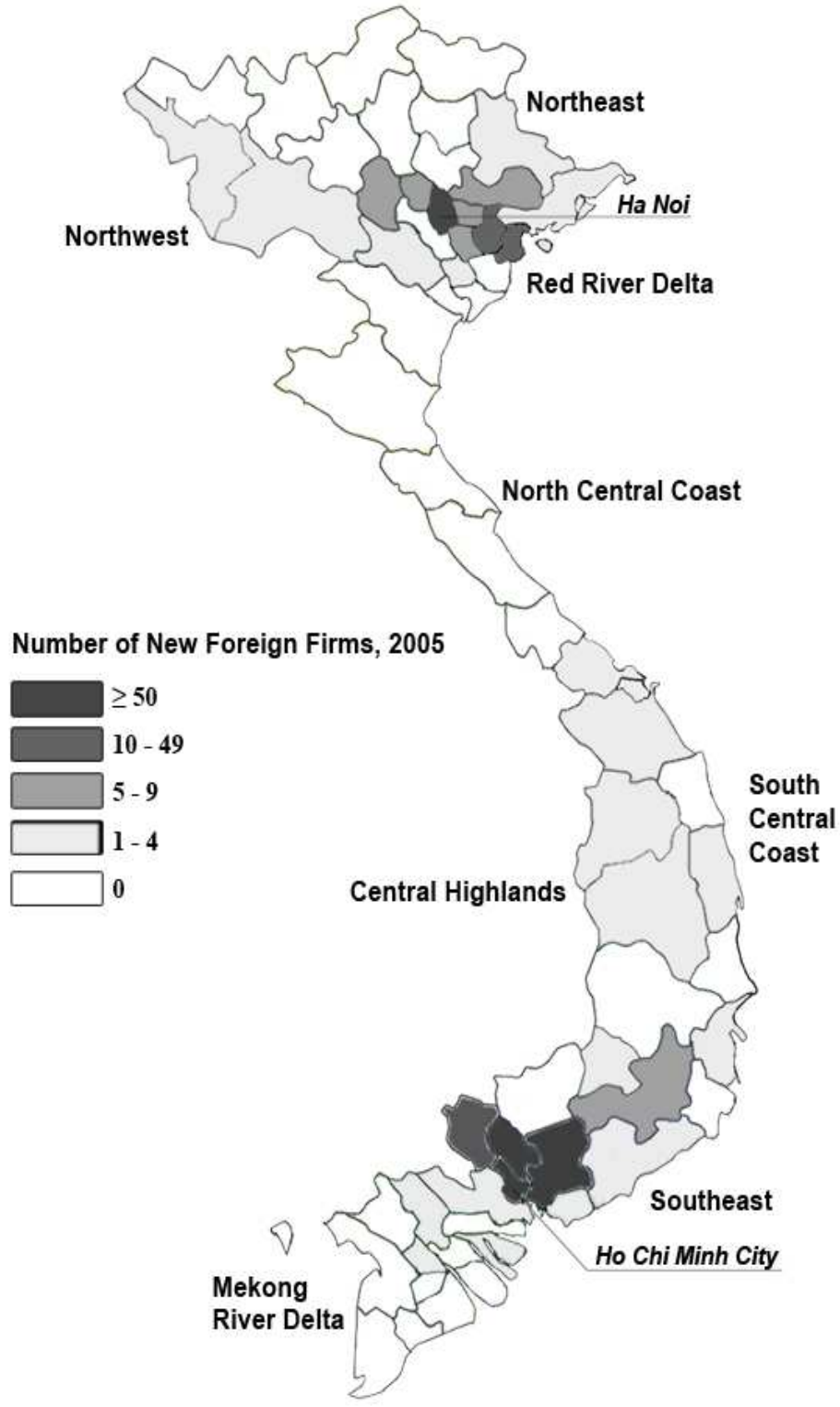
The dataset that is used in this study is obtained from the yearly surveys of the enterprises operating in Vietnam conducted by the General Statistics Office of Vietnam since 2000. These are comprehensive surveys covering all state enterprises, non-state enterprises that have equal or greater than 10 employees, 20% of sampled non-state enterprises with fewer than 10 employees, and all foreign enterprises across 64 provinces and cities in Vietnam. The contents of the surveys cover indicators to identify enterprises including their name, address, type, and economic activities of the enterprises, and indicators to reflect production situations of the enterprises such as their employees, income of employees, asset and capital source,

turnover, profit, contributions to the state budget, investment capital, taxes and other obligations to the government, job training, and evaluations on the investment environment. To our knowledge, this dataset has not been used for studies on location choices by foreign investors in Vietnam.

The sample includes foreign investors that started their activities in 2005. The newly created foreign firms in 2005 are identified by using tax codes that are unique for each firm to merge the cumulative number of foreign firms in 2005 with those in 2004, 2003, 2002, 2001 and 2000. Then the years in which foreign firms started their operation and industrial codes are used to track back the data to guarantee that the remaining firms are the newly created foreign firms in 2005. In sum, there were 568 new foreign firms created in 2005. The previous investors that are used to form the agglomerations are the cumulative number of foreign or Vietnamese firms up to 2004. In this study, firms from all industrial sectors in 4-digit industries and in all forms of ownership such as 100% foreign-owned and joint venture firms are included in the regression models.

Fig.3.2 depicts the geographical patterns of new foreign firms in 2005 by province. By looking at the color changes over the provinces on Fig. 1, we can see that most of the new foreign firms concentrated in Ho Chi Minh City and its two neighboring provinces, Binh Duong and Da Nang that belong to the Southeast region, and Hanoi that belongs to the Red River Delta region. While just these four provinces and cities accounted for 78.5% of the 568 new foreign firms in 2005, 30 out of the 64 provinces in Vietnam had no new foreign investors in 2005. Most of these provinces are in the North Central Coast, the Northwest and the Mekong River Delta regions (see Appendix 1 for more details).

Figure 2: The geographical distribution of new foreign firms in Vietnam, 2005



Source: Based on the dataset of the Survey on Enterprises in Vietnam in 2005, the GSO.

5. Methodology and empirical results

Various modeling approaches and levels of aggregation have been used for analyzing industrial location such as ordinary least squares (Boudier-Bensabaa, 2005), conditional logit model (Head *et al.*, 1995; Crozet *et al.*, 2004; Guimares and Figueiredo, 2000), negative binomial regression model (Meyer and Nguyen, 2005; Coughlin and Segev, 2000), and Generalized Method of Moments (Cheng and Kwan, 2000). These procedures have been applied to foreign direct investment aggregated to the country level or the provincial level and, more frequently in recent years, to the firm level. By virtue of possessing a large and detailed dataset, this study can use two different models to examine the three hypotheses: the negative binomial regression model and the conditional logit model. With the negative binomial regression model, it is possible to use only aggregated number of foreign firms at the provincial level. However, this model cannot exclude the fixed effects of the provinces that may lead to the biasness of the estimates. The conditional logit model can overcome this disadvantage by using the information about each foreign firm.

5.1. Agglomeration effects on location choices by foreign firms in Vietnam at the provincial level

The model and variables

Following the works of Coughlin and Segev (2000) and Meyer and Nguyen (2005), the negative binomial regression model is used with the provincial-level data across the sixty four provinces in Vietnam. A Poisson or a negative binomial distribution is frequently used to characterize processes that generate nonnegative integer outcomes such as the number of accidents that occur at a particular intersection. Thus, the number of new foreign firms locating in a specific province is a reasonable candidate for a Poisson or a negative binomial distribution. If there is overdispersion (i.e. the variance greater than the mean), estimates from the Poisson regression model will be inefficient (Long, 1997). In this case, the negative binomial regression model is preferred.

Dependent variables

The dependent variables are the number of newly created foreign firms and the number of new foreign firms by province that operate in the manufacturing sector. In 2005, there were 568 new foreign firms of which 381 were manufacturers. The Poisson or the negative binomial regression model only allows examining

Hypotheses 1 and 2. Table 2 and Table 3 present the descriptive statistics and the correlations of variables used in this analysis.

Agglomeration variables

In order to examine Hypothesis 1 that new foreign investors tend to locate in provinces where many other foreign firms have already existed, the cumulative number of foreign firms by province up to 2004 is used as a proxy. To examine Hypothesis 2 that firms in the same industries tend to cluster in particular regions, the cumulative number of foreign and Vietnamese firms in the manufacturing sector at provincial level up to 2004 is used as proxies. By 2004, there were 3145 foreign firms of which 2325 operate in the manufacturing sector, and 88420 Vietnamese firms of which 18125 are manufacturers.

Control variables

It is expected that provincial endowment factors can influence a firm's desire to invest in a particular province, such as the size of the provincial economy, the size of the provincial market, infrastructure, human resources, and geographical location. For instance, Ho Chi Minh City will always have a larger market than Ha Tinh province. Binh Duong will always enjoy a better location than Kon Tum or Ca Mau. Ha Noi will always have better infrastructure and more developed human resources than Ha Giang. So, the larger and more developed provinces such as Ho Chi Minh City, Ha Noi, Ba Ria – Vung Tau, Da Nang, Dong Nai, and Hai Phong will have more competitiveness simply because of their initial endowments. For this reason, following the work of Meyer and Nguyen (2005), the control variables that are included in the regression model are the size of local consumer market measured by the population of province, GDP by province, human capital development measured by the number of undergraduate students by province, and infrastructure conditions proxied by the number of industrial zones by province and the distance to the nearest big harbor. These data are cumulated up to 2004 and taken from the Statistical Yearbooks of Vietnam, the GSO.

Table 2: Descriptive statistics

Variables	Description	Mean	S.D.	Minimum	Maximum
1. New firm	Number of newly created foreign firms by province in 2005	8.87	30.34	0	201
2. New manufacturing firm	Number of newly created foreign manufacturing firms by province in 2005	5.95	18.20	0	109
3. Foreign firm	Number of foreign firms by province, cumulated up to 2004	49.14	157.45	0	1004
4. Foreign manufacturing firm	Number of foreign manufacturing firms by province, cumulated up to 2004	36.32	117.39	0	652
5. Vietnam manufacturing firm	Number of Vietnamese manufacturing firms by province, cumulated up to 2004	283.20	670.73	10	4845
6. Population	Average population, in thousands by province, in 2004	1281.74	865.72	295.1	5730.8
7. Student	Number of undergraduate students by province in 2004	21635.31	76338.09	356	498928
8. GDP	GDP in million VND by province in 2004	1.13e+07	2.07e+07	818111	1.37e+08
9. Industrial zone	Number of industrial zones by province cumulated up 2004	0.95	2.40	0	12
10. Distance to harbor	The distance in km to the nearest big harbor by province	149.99	99.26	0	387.61

Table 3: Correlations in the dataset

Variables	Notation	1	2	3	4	5	6	7	8	9	10
1. New firm	newfirm	1									
2. New manufacturing firm	newmanfirm	0.89	1								
3. Foreign firm	forfirm04	0.99	0.90	1							
4. Foreign manufacturing firm	manfirm04	0.95	0.97	0.97	1						
5. Vietnam manufacturing firm	manvn04	0.89	0.62	0.87	0.75	1					
6. Population	pop04	0.62	0.40	0.61	0.51	0.76	1				
7. Student	student04	0.65	0.40	0.64	0.48	0.84	0.59	1			
8. GDP	gdpmil04	0.74	0.49	0.74	0.63	0.84	0.68	0.66	1		
9. Industrial zone	iz04	0.83	0.84	0.86	0.88	0.66	0.48	0.42	0.71	1	
10. Distance to harbor	harbordis04	-0.33	-0.36	-0.34	-0.35	-0.31	-0.29	-0.20	-0.32	-0.38	1

Empirical Results

The empirical analysis is implemented as follows. First, Hypothesis 1 is examined to see if the number of already existing foreign firms in a province affects location decision by a new foreign in that province. Then, the regression model is applied to the foreign manufacturing firms for testing Hypothesis 2.

Table 4: Agglomeration effects in the negative binomial and Poisson models

Independent Variables	Negative binomial regression		Poisson regression	
	New firm 1	New mnf firm 2	New firm 3	New mnf firm 4
Foreign firm	0.0086** (0.0040)	-	0.0034**** (0.0005)	-
Foreign manufacturing firm	-	0.0140** (0.0071)	-	0.0059**** (0.0012)
Vietnam manufacturing firm	-	-0.0004 (0.0013)	-	-0.0010** (0.0004)
Population	-0.0004 (0.0004)	-0.0002 (0.0005)	0.0001 (0.0001)	0.0003** (0.0001)
Student	3.91e-06 (3.49e-06)	7.65e-06 (4.85e-06)	6.15e-06**** (4.33e-07)	7.63e-06**** (1.06e-06)
GDP	-2.14e-08 (1.79e-08)	-2.97e-08 (3.77e-08)	-3.39e-08**** (7.17e-09)	-1.10e-08 (1.10e-08)
Industrial zone	-0.0058 (0.1568)	-0.1180 (0.2089)	0.1591**** (0.0292)	0.0654 (0.0525)
Distance to harbor	-0.0074**** (0.0022)	-0.0082**** (0.0024)	-0.0083**** (0.0013)	-0.0101**** (0.0015)
α	1.4781 (0.4485)	1.5355 (0.4926)		
Obs (provinces)	61	61	61	61
Pseudo R2	0.18	0.17	0.86	0.80
Chi square	53.01****	46.29****	2036.72****	1192.52****

Note: Standard error in parentheses with significance at the **** 0.5%, *** 1%, **5%, and *10% levels.

New mnf firm: New manufacturing firm

After testing for Hypothesis Ho: $\alpha = 0$, we find a strong and statistically significant evidence of overdispersion [chibar2 (01) = 89.52, *p*-value < 0.01][3]. So the negative binomial regression model is used instead of the Poisson regression model to estimate empirical results. The number of observations is 61 because the information about the variable *student*, the number of undergraduate students cumulated up to 2004, is missing for three provinces - Lai Chau, Dac Nong, and Hau Giang – for three years of 2000, 2001 and 2002 because the Vietnamese government divided the 61 existing provinces into 64 in 2003.

The empirical results in Column 1 of Table 4 show evidence of agglomeration economies as the coefficient of the variable *foreign firm*, the cumulative number of foreign firms cumulated up to 2004, is positive and statistically significant. This result suggests that new foreign firms are more likely to locate in provinces with greater numbers of already existing foreign firms.

In order to test Hypothesis 2, the sample was restricted to include only newly created foreign firms in manufacturing sector. The negative binomial regression model was used since the testing of Hypothesis Ho: $\alpha = 0$ shows strong evidence of overdispersion [chibar2 (01) = 76.37, p-value < 0.01].

In Column 2 of Table 4, the positive and statistically significant coefficient of the variable *foreign manufacturing firm*, the number of foreign manufacturing firms cumulated up to 2004, supports the hypothesis that foreign firms in the same industries are likely to locate near each other. However, the negative and statistically insignificant estimate of the variable *Vietnam manufacturing firm*, the number of Vietnamese manufacturing firms cumulated up to 2004, suggests that the locations of Vietnamese firms do not influence the location decisions by foreign firms in the same industries.

Different from the results of Meyer and Nguyen (2005), most of the control variables are statistically insignificant except the variable *distance to harbor*[4], the distance to the nearest big harbor. The negative sign of the variable *distance to harbor* means that the nearer a province is to a big harbor, the more attractive it is to

[3] The Poisson regression model accounts for only observed heterogeneity (i.e., observed difference among sample members). In practice, the Poisson regression model rarely fits due to *overdispersion*. That is, the model underfits the amount of dispersion in the outcome, leading to biased-downward standard errors that result in spuriously large *z*-values and spuriously small *p*-values. The negative binomial regression model addresses the failure of the Poisson regression model by adding a parameter, α, that determines the degree of dispersion in the predictions by reflecting unobserved heterogeneity among observations (see Long and Freese, 2006 for more details).

[4] The study has run the regression model with the quadratic variable *harbordissq* (the square value of the variable *distance to harbor*) and found that the coefficient of *harbordissq* is statistically insignificant while the coefficient of the variable *distance to harbor* is still statistically significant and negative although the significance is reduced. This evidence suggests that the effect of *distance to harbor* on the location decisions by foreign firms is linear.

foreign investors. This evidence suggests that foreign investors prefer to locate in a place with upgraded infrastructure to reduce transportation costs.

Columns 3 and 4 of Table 4 present the estimates of the Poisson regression model. By contrast with the results of the negative binomial regression model, the coefficients of most variables are highly statistically significant and the Pseudo R2 is very high. The reason is that the Poisson regression model in this case ignores unobserved heterogeneity among observations, leading to biased-downward standard errors that result in spuriously large *z*-values and spuriously small *p*-values.

It is noted that Table 3 shows high correlations between the variables *foreign manufacturing firms* and *Vietnamese manufacturing firms* as well as between the variables *population* and *GDP*. We suspect that the result of non-significance of the variable related to the presence of Vietnamese firms is due to collinearity problems among explanatory variables. In order to check if the empirical results suffer from these problems, we have re-run some additional regressions inserting alternatively the variable *foreign manufacturing firms* and the variable *Vietnamese manufacturing firms* as well as between the variable *population* and the variable *GDP* and find that the estimated results are robust and do not appear to result from collinearity among the regressors (See Appendix 2 for more details).

Overall, the regression results support the hypotheses that foreign firms agglomerate. Foreign firms in Vietnam are likely to locate near each other and near other foreign firms in the same industries. However, the locations of Vietnamese firms have no influence on the location decisions by foreign firms in the same industries. The findings are consistent with many previous studies on location choices by foreign investors in different countries such as the studies of Boudier-Bensebaa (2005), Meyer and Nguyen (2005), Head *et al.* (1995), Cheng and Kwan (2000), and Crozet *et al.* (2004).

5.2. Agglomeration effects on location choices by foreign investors in Vietnam at the firm level

By using the negative binomial model, we find the evidence of agglomeration effects. However, the concern is that there may be provincial fixed effects which generate a misleading correlation between the cumulative number of firms which have entered a province and the new entries in the year in question. These results may be caused by unobserved heterogeneity across provinces leading to a spurious agglomeration coefficient. Suppose that we have attributed the entry to clustering while it is in fact the better facilities of a province that are responsible. These facilities are defined as fixed effects if they are unchanged over time, unobservable

and affect the number of new entries in provinces. If unobserved effects correlate with the explanatory variables, the estimation will be biased and inconsistent.

In order to eliminate fixed effects of the provinces, the conditional logit model is used since this model bases on the information about individual firms to estimate the effects of agglomeration on its location choice. With the detailed and precise information about each foreign firm operating in Vietnam, it is feasible to apply this model to examine all the three hypotheses mentioned in section 3.

The conditional logit model is widely used in previous empirical works on agglomeration effects (Head *et al.*, 1995; Crozet *et al.*, 2004; Shaver and Flyer, 2000; Guimaraes *et al.*, 2000). This model is derived from the result of McFadden (1974) with the assumption that each investor chooses a location that will yield the highest profit. Profit depends on the available inputs that go into firms' production function including agglomeration effects stemming from economic activities of near similar firms. In this model, the information about the location choice that an investor made and attributes for the chosen location and other locations in the choice set are exploited.

Following Head *et al.* (1995), the study considers that the investor *i*, if it locates in province *j*, will derive an expected profit of Π_{ij}. This investor chooses the location with the greatest expected profitability that can be represented as followed:

$$\Pi_{ij} = \alpha_j + \beta' X_{ij} + \varepsilon_{ij}$$

where α_j includes the characteristics of province *j*. α_j is considered as province-specific endowment effects that determine the attractiveness of provinces to investors[5]. X_{ij} is agglomeration variables measured as the count number of firms cumulated up to 2004. Each measure varies across investors *i*, because investors differ by industry and country of origin. ε_{ij} is an investment location specific random disturbance that is attributable to errors associated with imperfect perception and optimization by decision makers and unobservable location characteristics that affect the profitability of locating in a given site.

The investor *i* prefers the location *j* among the choice set *M* if it yields higher profits than any other possible choices:

$$\Pi_{ij} > \Pi_{ik} \quad \forall k, k \neq j, \text{ and } j, k \in M.$$

The probability of choosing the location *j* is thus:

$$\text{Prob}(\Pi_{ij} > \Pi_{ik}) \quad \forall k, k \neq j.$$

[5] Head *et al.* (1995) show that in both theories of localization, *endowment-driven localization and agglomeration model of industry localization*, firms in the same industry cluster geographically. However, only in the presence of agglomeration externalities does the clustering *add* to the attractiveness of the location.

McFadden (1974) shows that if, and only if, ε_{ij} is distributed as a Type I Extreme Value independent random variable, then the probability that a location *j* yields the highest profitability for investor *i* among all the alternative locations in the choice set *M* is presented by the logit model:

$$\Pr(ij) = \frac{\exp(\alpha_j + \beta' X_{ij})}{\sum_M \exp(\alpha_m + \beta' X_{im})} \qquad j, m \in M$$

The maximum likelihood techniques are used to estimate endowment effects and agglomeration effects.

Variables

As the part using the negative binomial negative model, the data in this part is from the surveys of all firms operating in Vietnam conducted by the General Statistics Office of Vietnam since 2000. In the conditional logit model, the information about the industry, the country of origin, and the location of each foreign firm is used. The attributes of provinces in the location choice set are collected from the Statistical Yearbooks of Vietnam. Table 5 and Table 6 present the descriptive statistics and the correlations of variables used in this model.

Dependent variable

The dependent variable is the province chosen by each foreign firm that was newly created in 2005. In total, there were 568 new foreign firms that distribute in 34 provinces among 64 provinces in Vietnam. Conditional logit model requires that all choices be selected at least once. So, 30 provinces that are not selected any time from the choice set are removed, including Ha Tay, Nam Dinh, Ninh Binh, Ha Giang, Cao Bang, Lao Cai, Bac Kan, Tuyen Quang, Yen Bai, Thai Nguyen, Lai Chau, Thanh Hoa, Nghe An, Ha Tinh, Quang Binh, Quang Tri, Quang Ngai, Phu Yen, Dak Lak, Ninh Thuan, Binh Phuoc, An Giang, Tien Giang, Vinh Long, Kien Giang, Hau Giang, Tra Vinh, Soc Trang, Bac Lieu, and Ca Mau. Most of these provinces are from the Northeast, the North Central Coast, and the Mekong River Delta regions. The other 34 provinces create a set of unordered choice for each foreign firm, say, $M = 1, 2, \ldots, 34$. Let y_{ij} ($j \in M$) be a dependent variable for the choice actually chosen by the i^{th} foreign firm. That is, $y_{ij} = 1$ if foreign firm *i* chooses the location *j*, and $y_{ij'} = 0$ for $j' \neq j$; $j, j' \in M$. In total, there are 19312 observations that equal 568 foreign firms multiplied with 34 provinces.

Agglomeration variables

The study estimates the effects of three types of agglomerations on the location choices by foreign investors in Vietnam. In each case, the agglomeration is measured as cumulative counts of firms up to 2004. It is noted that cumulated up to 2004, there were 3145 foreign firms and 88420 Vietnamese firms. Following the work of Guimaraes *et al.* (2000), Head *et al.* (1995) and Crozet *et al.* (2004), there are three types of agglomeration effects as follows:

- Foreign-specific agglomeration: the cumulative number of foreign firms by province up to 2004 is used as a proxy.
- Industry-specific agglomeration: the cumulative number of Vietnamese firms in the same 4-digit industries by province, the cumulative number of foreign firms in the same 4-digit industries by province and the cumulative number of foreign firms in the same industries in the neighboring provinces up to 2004 are used as proxies.
- Country-specific agglomeration: the cumulative number of foreign firms from the same countries of origin by province up to 2004 is used as a proxy.

Including the cumulative number of Vietnamese firms in the same 4-digit industries is a strategy to separate agglomeration and endowment effects. The reason is that although α_j captures the attractiveness of province j to the "average" investors, unobserved characteristics of investors can make some provinces more attractive to certain investors. For example, a firm in an industry with high factor intensities will choose provinces with abundant endowments of these factors. This suggests that industry-level agglomeration variables will be correlated with the unobserved factor conditions pertaining to that industry that constitute the error term in the model. This problem can be solved by including province- and industry-specific characteristics. However, this strategy is infeasible with the sample of 568 foreign firms in about 155 different 4-digit industries. The significant attraction of the old firms to new ones in the same industries or the countries of origin, after controlling for the patterns of Vietnamese firms, can provide the evidence of agglomeration effects (see Head *et al.*, 1995). In other words, the number of Vietnamese firms in the same 4-digit industries acts as a proxy for industry-specific endowment effects.

Using the idea of Head *et al.* (1995), the number of foreign firms in the neighboring provinces is included in the model. This variable allows the possibility that, for example, Binh Duong province is attractive to wearing apparel manufacturers not only because of the wearing apparel producers there but also

because of the wearing apparel producers in the neighboring provinces: Ho Chi Minh City, Tay Ninh, Dong Nai, Ba Ria-Vung Tau, Long An, and Tien Giang.

Control variables

In the conditional logit model, the same control variables of the negative binomial regression model are used. These control variables reflect the characteristics of the provinces that are considered as province-specific endowment effects determining the attractiveness of the provinces to foreign investors. The control variables for the size of local consumer market measured by the population of province, GDP by province, human capital development measured by the number of undergraduate students by province, and infrastructure conditions proxied by number of industrial zones by province and the distance to the nearest big harbor are included in the model. These data are cumulated up to 2004 and taken from the Statistical Yearbooks of Vietnam, the GSO.

Table 5: Descriptive statistics

Variables	Obs	Description	Mean	S.D.	Min	Max
1. Choice	19312	Dummy variable which equals 1 if firm i chooses location j and equals 0 for other location j', j≠ j' and j, j' belong to the location choice set	0.03	0.16	0	1
2. Foreign firm	19312	The cumulative number of foreign firms by province up to 2004	89.29	206.10	0	1004
3. Vietnamese firm	19312	The cumulative number of Vietnamese firms in the same 4-digit industries by province up to 2004	14.48	65.74	0	1905
4. Same industry	19312	The cumulative number of foreign firms in the same 4-digit industries by province up to 2004	2.00	9.32	0	146
5. Neighboring firm	19312	The cumulative number of foreign firms in the same 4-digit industries in neighboring provinces up to 2004	8.43	23.13	0	201
6. Same country	18802*	The cumulative number of foreign firms from the same countries of origin by province up to 2004	12.13	41.67	0	328
7. Population	19312	Average population in thousands by province in 2004	1344.40	922.07	366.1	5730.8
8. Student	18744**	Number of undergraduate students by province in 2004	35782.88	100522.5	434	498928
9. GDP	19312	GDP in million VND by province in 2004	1.57e+07	2.72e+07	1527060	1.37e+08
10. Industrial zone	19312	Number of industrial zones by province in 2004	1.64	3.08	0	12
11. Distance to harbor	19312	The distance in km to the nearest big harbors by province	115.07	94.90	0	384.42

Notes: * In 568 new foreign firms in 2005, there are 15 firms without information about countries of original (18802 obs = 19312 – 15x34).

** The information about number of students is missing in one province of the location choice set (18744 obs = 19312 – 1x568).

Table 6: Correlations in the dataset

Variables	Notation	1	2	3	4	5	6	7	8	9	10	11
1. Choice	choice	1										
2. Foreign firm	forfirm04	0.41	1									
3. Vietnamese firm	vnfirm4dgsic	0.25	0.47	1								
4. Same industry	same4dgsic	0.34	0.53	0.59	1							
5. Neighboring firm	border4dgsic	0.07	0.28	0.13	0.37	1						
6. Same country	samecountry	0.32	0.68	0.31	0.42	0.26	1					
7. Population	pop04	0.33	0.78	0.51	0.44	0.12	0.48	1				
8. Student	student04	0.26	0.62	0.49	0.32	0.00	0.34	0.73	1			
9. GDP	gdpmil04	0.30	0.74	0.46	0.41	0.23	0.45	0.77	0.65	1		
10. Industrial zone	iz04	0.34	0.86	0.33	0.44	0.40	0.61	0.58	0.39	0.71	1	
11. Distance to harbor	harbordis04	-0.13	-0.34	-0.13	-0.18	-0.21	-0.24	-0.34	-0.17	-0.30	-0.40	1

Empirical results

Table 7 presents the agglomeration coefficients generated by maximum likelihood estimation. The highly statistically significant coefficients of the variables *foreign firm*, the cumulative number of foreign firms by province up to 2004 and *Vietnamese firm*, the cumulative number of Vietnamese firms in the same 4-digit industries by province up to 2004, in Column 1 reveal that new foreign firms are likely to locate in provinces where already existed a relatively large number of foreign firms in the same industries.

In Columns 2, the cumulative number of foreign firms in the same 4-digit industries up to 2004 (*same industry*) is added to the regression model. The positive and highly statistically significant coefficient of the variable *same industry* proves that the locations of new foreign investments are influenced by the previous location choices by other foreign firms in the same industries. Head *et al.* (1995) consider this phenomenon as the "follow the leader" pattern of foreign firms; that is difficult to interpret as anything other than agglomeration effects.

However, when we insert the variable related to the number of foreign firms in the same industry (*same industry)*, the coefficient of the cumulative number of Vietnamese firms in the same 4-digit industries (*Vietnamese firm*) becomes negative and statistically insignificant while there is no change for the variable *foreign firm*. This result shows that the positive correlation (0.60) between *same industry* and *Vietnamese firm* is important. Vietnamese firms and foreign firms in the same industries tend to invest in the same locations. If we do not include the variable *same industry* in the regression, its effect is attributed to *Vietnamese firm* giving a positive bias to the *Vietnamese firm* coefficient. Whenever we include *same industry* variable, the coefficient of *Vietnamese firm* is negative and insignificantly different from zero. Moreover, by running the likelihood ratio tests we find that the models which omit the variable *same industry* appear misspecified and are dominated by the models including it in the regressions.

Compared with Head *et al.* (1995), this result reflects a different tendency in the location decisions by foreign investors in Vietnam from that of Japanese investors in the United States. Head *et al.* (1995) found that Japanese firms prefer to locate near US firms in the same industries. The regression model, however, shows that the location choices by new foreign investors are not influenced by the locations of Vietnamese firms. Different from the location patterns of US and Japanese firms, Appendix 1 shows that the location distributions of foreign firms and Vietnamese firms are not very matched. While most foreign investments concentrate in the Red River Delta and Southeast regions, especially in the cities and provinces of Hanoi, Ho Chi Minh City,

Binh Duong, and Dong Nai, Vietnamese firms are distributed quite evenly in all provinces. The negative and statistically insignificant coefficient of the variable *Vietnamese firm* encourages us to believe that the estimates of agglomerations are not influenced by industry-specific endowment effects.

Table 7: Agglomeration effects in the conditional logit model

Independent variable	Dependent variables: location choice 1	2	3	4
Foreign firm	0.0042**** (0.0006)	0.0038**** (0.0007)	0.0039**** (0.0006)	0.0033**** (0.0006)
Vietnamese firm	0.0015**** (0.0004)	-0.0005 (0.0005)	-0.0004 (0.0004)	-0.0004 (0.0004)
Same industry	-	0.0226**** (0.0032)	0.0207**** (0.0031)	0.0195**** (0.0031)
Neighboring firm	-		-0.0073*** (0.0026)	-0.0081**** (0.0026)
Same country	-		-	0.0032**** (0.0008)
Population	0.0006*** (0.0002)	0.0007*** (0.0003)	0.0006*** (0.0002)	0.0006*** (0.0002)
Student	4.50e-06**** (4.48e-07)	4.98e-06**** (4.54e-07)	4.86e-06**** (4.56e-07)	4.91e-06**** (4.56e-07)
GDP	-5.08e-08**** (1.12e-08)	-5.14e-08**** (1.12e-08)	-5.28e-08**** (1.13e-08)	-5.18e-08**** (1.12e-08)
Industrial zone	0.1078**** (0.0323)	0.1081**** (0.0324)	0.1225**** (0.0328)	0.1263**** (0.0328)
Distance to harbor	-0.0037**** (0.0012)	-0.0037**** (0.0012)	-0.0037**** (0.0012)	-0.0037**** (0.0012)
Log-likelihood	-1203.2	-1175.21	-1171.4	-1163.8
Pseudo R2	0.37	0.39	0.39	0.40
Chi square	1453.8****	1509.7****	1517.4****	1532.3****
No. of choosers	568	568	568	568
No. of choices	34	34	34	34

Note: Standard error in parentheses with significance at the **** 0.5%, *** 1%, ** 5%, and * 10% levels.

The negative and statistically significant coefficient of the variable *neighboring firm* in Columns 3 and 4 indicates that a larger number of foreign firms in the same

industries in a province decrease the attractiveness of its neighboring provinces to new foreign investors. It appears that there is competition among provinces in attracting foreign investors. In Column 4, the number of foreign firms from the same countries of origin is added in the regression model to determine whether firms from the same countries of origin tend to locate near each other. The positive and statistically significant coefficient of the variable *same country*, the cumulative number of foreign firms from the same countries of origin up to 2004, indicates that new foreign firms benefit from locating near firms from the same countries of origin. The larger coefficient of the variable *same industry* than that of the variable *same country* suggests that the benefits foreign firms gain from industry-specific agglomerations are higher than from country-specific agglomerations.

Different from the results of the negative binomial model, all control variables here are statistically significant except the negative sign of the variable GDP is out of expectation. These results indicate that the characteristics of the provinces are important determinants in attracting foreign investors.

As discussed in the previous part, we are also concerned that high correlations between the variables *same industry* and *Vietnamese firms* as well as between the variables *population* and *GDP* may lead to the result of non-significance of the variable related to the presence of Vietnamese firms in the same industry. In order to check if the empirical results suffer from collinearity problems, we have re-run some additional regressions inserting alternatively the variable *same industry* and the variable *Vietnamese firms* as well as between the variable *population* and the variable *GDP* and find that the estimated results are robust and do not appear to result from collinearity amongst the regressors. (See Appendix 2 for more details).

In summary, the empirical results support the hypotheses that foreign investors are not only likely to locate near other foreign firms but also prefer to locate near foreign firms in the same industries and from the same countries of origin due to the benefits from agglomeration economies. Moreover, we found that provinces in Vietnam compete with each other to attract foreign firms and location choices by foreign investors are not affected by location of domestic firms.

5.3. Robustness tests

In order to investigate whether the empirical results are robust, the both regression models are re-estimated by using a variety of sub-samples of the dataset. Following Guimaraes *et al.* (2000), it is possible to test the existence of agglomeration economies

in location decisions by foreign investors according to firms' capital ownership and size.

Table 8: Agglomeration effects in the negative binomial regression model

Independent Variables	nffewer100emp 1	nf100%forcap 2	mffewer100emp 3	mf100%forcap 4
Foreign firm	0.0074** (0.0039)	0.0086** (0.0045)	-	-
Foreign manufacturing firm	-	-	0.0144* (0.0081)	0.0160** (0.0082)
Vietnam manufacturing firm	-	-	-0.0009 (0.0014)	-0.0006 (0.0014)
Population	-0.0003 (0.0004)	-0.0004 (0.0005)	-0.0001 (0.0005)	-0.0001 (0.0006)
Student	4.67e-06 (3.45e-06)	3.56e-06 (3.85e-06)	9.57e-06 (5.17e-06)	8.73e-06 (5.27e-06)
GDP	-1.81e-08 (1.76e-08)	-2.09e-08 (1.92e-08)	-1.96e-08 (3.82e-08)	-3.16e-08 (4.63e-08)
Industrial zone	0.0263 (0.1575)	0.0017 (0.1761)	-0.1404 (0.2425)	-0.1696 (0.2360)
Distance to harbor	-0.0080**** (0.0025)	-0.0087**** (0.0026)	-0.0104**** (0.0032)	-0.0108**** (0.0030)
Obs	61	61	61	61
Pseudo R2	0.20	0.17	0.19	0.18
Chi square	51.68****	46.93****	42.58****	42.94****

Notes: Standard error in parentheses with significance at the **** 0.5%, *** 1%, ** 5%, and * 10% levels.

nffewer100emp: new firms have fewer than 100 employees

nf100%forcap: new firms of 100% foreign capital

mffewer100emp: new firms have fewer than 100 employees in manufacturing sector

mf100%forcap: new firms of 100% foreign capital in manufacturing sector

In the previous parts, all kinds of investments with foreign participations i.e., 100% foreign capital owned firms and joint venture enterprises are included in the regression models. For the first test of the results' robustness, only newly created firms of 100% foreign capital are used. We argue that these firms can decide the locations by themselves while the decisions by join venture enterprises somehow depend on the both Vietnamese and foreign sides. Of 568 newly created foreign firms in 2005, there were

491 firms of 100% foreign capital, of which 347 are operating in the manufacturing sector.

Table 9: Agglomeration effects in the conditional logit model

Independent Variables	Dependent variable: location choice		
	nffewer100emp	nfmore100emp	nf100%forcap
	1	2	3
Foreign firm	0.0030****	0.0033***	0.0033****
	(0.0008)	(0.0012)	(0.0007)
Vietnamese firm	0.0000	-0.0002	-0.0003
	(0.0005)	(0.0013)	(0.0004)
Same industry	0.0317****	0.0149**	0.0193****
	(0.0046)	(0.0067)	(0.0032)
Neighboring firm	-0.0083**	-0.0024	-0.0074***
	(0.0041)	(0.0032)	(0.0027)
Same country	0.0032****	0.0032**	0.0023***
	(0.0010)	(0.0015)	(0.0008)
Population	0.0006**	0.0006	0.0006**
	(0.0003)	(0.0004)	(0.0002)
Student	5.93e-06****	-9.41e-08	4.85e-06****
	(5.13e-07)	(1.80e-06)	(5.19e-07)
GDP	-5.31e-08****	-4.24e-08**	-5.11e-08****
	(1.39e-08)	(1.93e-08)	(1.12e-08)
Industrial zone	0.1473****	0.0627	0.1457****
	(0.0400)	(0.0602)	(0.0347)
Distance to harbor	-0.0042***	-0.0031*	-0.0044****
	(0.0015)	(0.0019)	(0.0013)
Log-likelihood	-820.4	-302.6	-990.6
Pseudo R2	0.46	0.29	0.41
Chi square	1373.1****	240.8***	1361.4****
No. of choosers	445	123	491
No. of choices	34	34	34

Note: Standard error in parentheses with significance at the **** 0.5%, *** 1%, ** 5%, and *10% levels

nffewer100em: new firms have fewer than 100 employees in manufacturing sector.

nfmore100emp: new firms have equal or more than 100 employees in manufacturing sector.

To investigate how agglomeration economies affect location decisions by firms with different size, we divide new foreign firms created in 2005 into kinds: large and small ones. Foreign firms are defined small if they have fewer than 100 employees, otherwise they are considered large. It is argued that regions in general compete for

large firms. However, location is not a big concern for a giant firm because in any places it might have higher competitiveness than the others. In 2005, there were 445 new foreign firms with fewer than 100 employees, of which 265 are manufacturers. To make it more simple, we include only small foreign firms in the negative binomial model, but include both of small and large firms in the conditional logit model.

The empirical results of the negative binomial regression and conditional logit models with the restricted samples are presented in Tables 3.8 and 3.9. Despite the smaller dimensions of the samples, the coefficients of variables are remarkably stable. All the agglomeration variables that were statistically significant in Tables 3.4 and 3.7 are still statistically significant in these regressions (see Table 8 and Columns 1 and 3 of Table 9).

However, the double coefficient of the variable *same industry*, the cumulative number of foreign firms in the same 4-digit industries up to 2004, in Column 1 compared with that of Column 2 (Table 9) shows that small foreign firms have a stronger motivation to locate near other foreign firms in the same industries than large foreign firms. This seems consistent with the argument of Shaver and Flyer (2000) that under the existence of agglomeration economies, small firms will have greater benefits since the agglomeration externalities allow them to access technologies of near larger competitors.

By contrast with Shaver and Flyer (2000), large foreign firms in this study also agglomerate. However, the statistically insignificant coefficient of the variable *neighboring firm*, the cumulative number of foreign firms in the same 4-digit industries in neighboring provinces, shows that large firms do not care about the existence of firms in the same industries in the bordering provinces. Different from the estimation results of small foreign firms or total foreign firms, most control variables for the large foreign firms are statistically insignificant (see Column 2 of Table 9). It seems that the characteristics of provinces are not a big concern for a large foreign firm.

6. Conclusions

This study argues that agglomeration externalities influence the location decisions by foreign firms. The empirical results show that the location choices by new foreign firms in Vietnam are affected by the locations of the prior foreign investments in general and by those of firms in the same industries and from the same countries of origin in particular. These findings hold even when province-specific endowment and

industry-specific endowment effects are controlled by using the variables indicating the characteristics of each province and the industry-level stocks of Vietnamese firms. Moreover, we find that the geographical distributions of Vietnamese firms have no effect on the location choices by foreign investors and there is competition among provinces in attracting foreign investors. It is noted that the empirical results hold when we test the existence of agglomeration economies in location choices by foreign firms regarding their ownership and size.

These findings are consistent with the empirical results that are estimated for foreign investments in developed countries such as the United States, France, and Portugal (Head *et al.*, 1995; Crozet *et al.*, 2004; Guimaraes *et al.*, 2000). It indicates that the behavior by foreign investors in both developed and developing countries are probably similar. Their same motivations are to obtain the highest benefits when investing abroad. Apparently, the positive externalities such as technological spillovers will induce foreign firms to cluster in a particular region. Moreover, locating near each other creates a network of foreign firms that allows a foreign firm to access suppliers and to exchange information more easily. This network may consist of foreign firms in the same industries that are considered as industrial or vertical groups. These groups might be headed by large manufacturing companies whose members are component suppliers. Vertical linkages can create a pool of specialized intermediate inputs to an industry in greater variety and at lower cost as suggested by Marshall (1920). So, for example, a firm that produces plastic auto parts might be attracted to a province that has considerable auto production even if there is no concentration of plastic parts producers in that province (Head *et al.*, 1995).

This research contributes to the literature on agglomeration economies, location and foreign direct investment in some aspects. To the best of our knowledge, the study on location decisions by individual firms has never been carried out in Vietnam due to the lack of detailed data at firm level. This is also one of a very few studies of agglomeration effects on location choices by foreign investors in developing and transition economies. The empirical findings on agglomeration economies may be useful for provincial authorities in designing policies to attract more foreign direct investment. Benefits of agglomeration externalities suggest that authorities should create policies to draw *initial* investments into concentrated production regions such as industrial zones. Then the cumulative number of foreign firms will create positive agglomeration externalities and make that region more attractive. This policy has been implemented effectively in the small province Binh Duong in the Southeast region of Vietnam. In 2005, Binh Duong province accounted for 19.8% of the total foreign

investment in Vietnam while hosting only 2% of the total number of Vietnamese firms. This success is partially based on the policies of this province to establish many industrial zones and to create a good business environment for foreign investors from the first days when the central government granted the provinces more autonomy in the management of foreign investment.

This study has two limitations. The first is that the empirical results refer to only 2005. In order to see whether the results apply to other time periods, future research will have to work with larger dataset covering more years, so as to increase the cross time variance in the set of agglomeration variables. Moreover, there is a concern that as in the conditional logit model the observations related to provinces that were not selected by new foreign firms in 2005 are lost. This might potentially distort results if the cumulated number of foreign firms up to 2004 in these "omitted provinces" that used as a proxy for agglomeration effects is not trivial. By calculating this proxy, we find that the cumulated number of foreign firms up to 2004 in these "omitted provinces" accounted for a very small proportional, around 0.035% of the total number of foreign firms up to 2004. Our choice set of location thus may reinforce the results: those provinces there were not selected in the year 2005 are probably also provinces where the cumulated number of firms is negligible thus confirming the argument of agglomeration economies. Therefore, by working with larger dataset covering more years, we also can have more exact conclusions about agglomeration effects. The second limitation is that we have studied the location decisions by foreign firms only at the provincial level. The conditional logit model may work better with a smaller choice set. Therefore, future research should extend to macro areas by looking at the location choices by foreign firms at the regional level.

Appendix 1: The location distributions of firms in Vietnam

Region/ Province/ City	No. of newly created foreign firms in 2005	No. of cumulative foreign firms up to 2004	No. of cumulative Vietnamese firms up to 2004
Red River Delta	**128**	**650**	**24537**
Ha Noi	72	379	14698
Hai Phong	22	127	2498
Vinh Phuc	7	29	680
Ha Tay	0	24	1236
Bac Ninh	7	10	877
Hai Duong	10	42	1081
Hung Yen	7	26	526
Ha Nam	2	1	438
Nam Dinh	0	4	986
Thai Binh	1	6	851
Ninh Binh	0	2	666
Northeast	**15**	**99**	**6097**
Ha Giang	0	0	271
Cao Bang	0	1	262
Lao Cai	0	8	517
Bac Kan	0	1	242
Lang Son	2	10	324
Tuyen Quang	0	0	299
Yen Bai	0	4	356
Thai Nguyen	0	11	791
Phu Tho	6	24	966
Bac Giang	5	13	894
Quang Ninh	2	27	1175
Northwest	**3**	**9**	**1035**
Lai Chau	0	0	129
Dien Bien	1	0	251
Son La	1	2	272
Hoa Binh	1	7	383
North Centra Coast	**1**	**30**	**5343**
Thanh Hoa	0	7	1184
Nghe An	0	7	1422
Ha Tinh	0	2	547
Quang Binh	0	1	749
Quang Tri	0	3	478
Thua Thien - Hue	1	10	963
South Central Coast	**8**	**95**	**6167**
Da Nang	2	30	1908
Quang Nam	2	12	622

Quang Ngai	0	2	669
Binh Dinh	1	9	1031
Phu Yen	0	8	474
Khanh Hoa	3	34	1463
Central Highlands	**11**	**51**	**2829**
Kon Tum	1	0	253
Gia Lai	1	2	671
Dak Lak	0	1	832
Dak Nong	1	3	156
Lam Dong	8	45	917
Southeast	**396**	**2129**	**29737**
Ho Chi Minh	201	1004	22723
Ninh Thuan	0	4	329
Binh Phuoc	0	3	472
Tay Ninh	20	49	675
Binh Duong	111	625	1734
Dong Nai	62	373	2063
Binh Thuan	1	14	676
Ba Ria - Vung Tau	1	57	1065
Mekong Delta River	**6**	**82**	**12675**
Long An	2	48	1083
Dong Thap	1	2	966
An Giang	0	3	1139
Tien Giang	0	5	1489
Vinh Long	0	3	833
Ben Tre	1	3	964
Kien Giang	0	2	1759
Can Tho	2	13	1284
Hau Giang	0	0	338
Tra Vinh	0	0	446
Soc Trang	0	0	740
Bac Lieu	0	2	546
Ca Mau	0	1	1088
Total	**568**	**3145**	**88420**

Source: The GSO, the Enterprise Surveys in Vietnam in 2004 and 2005

Appendix 2: Robustness checks of the models

A. Negative binomial regression

As discussed in the previous parts, we are concerned that empirical results may suffer from collinearity problems among explanatory variables. In order to check robustness of the models, we have re-run some additional regressions inserting alternatively the variable *foreign manufacturing firms* (*manfirm04*) and the variable *Vietnamese manufacturing firms* (*manvn04*) as well as between the variable *population* (*pop04*) and the variable *GDP* (*gdpmil04*). We always include in the regressions all control variables, except the variable *pop04* and find that the estimated results are robust and do not appear to result from collinearity amongst the regressors, therefore confirming that foreign firms are likely to locate near each other and near other foreign firms in the same industries. However, their location choices are not affected by location of Vietnamese firms. It is noted that the conclusions are the same if we always include in the regressions all control variables, except the variable *gdpmil04.*

A1- Include only *forfirm04* and exclude *pop04*

```
Negative binomial regression                    Number of obs   =         61
                                                LR chi2(5)      =      52.05
Dispersion     = mean                           Prob > chi2     =     0.0000
Log likelihood = -119.7585                      Pseudo R2       =     0.1785

------------------------------------------------------------------------------
     newfirm |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
   forfirm04 |    .007203   .0034923     2.06   0.039     .0003583    .0140478
   student04 |   3.43e-06   3.45e-06     1.00   0.320    -3.32e-06    .0000102
    gdpmil04 |  -2.57e-08   1.78e-08    -1.44   0.149    -6.05e-08    9.22e-09
        iz04 |   .0233513   .1584438     0.15   0.883    -.2871929    .3338954
   harbordis |  -.0070175   .0022052    -3.18   0.001    -.0113396   -.0026954
       _cons |   1.525112   .3949372     3.86   0.000     .7510495    2.299175
-------------+----------------------------------------------------------------
    /lnalpha |    .395117   .3060559                     -.2047416    .9949755
-------------+----------------------------------------------------------------
       alpha |   1.484558   .4543577                      .8148579    2.704658
------------------------------------------------------------------------------
Likelihood-ratio test of alpha=0:  chibar2(01) =   89.58 Prob>=chibar2 = 0.000
```

A2- Include only *manfirm04* and exclude *pop04*

```
Negative binomial regression                    Number of obs   =         61
                                                LR chi2(5)      =      45.95
Dispersion     = mean                           Prob > chi2     =     0.0000
Log likelihood = -109.35944                     Pseudo R2       =     0.1736

------------------------------------------------------------------------------
  newmanfirm |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
   manfirm04 |   .0127158    .007093     1.79   0.073    -.0011861    .0266177
   student04 |   6.58e-06   3.67e-06     1.79   0.073    -6.24e-07    .0000138
    gdpmil04 |  -4.62e-08   3.00e-08    -1.54   0.123    -1.05e-07    1.26e-08
        iz04 |  -.0838512    .201955    -0.42   0.678    -.4796757    .3119733
   harbordis |  -.0080103   .0024052    -3.33   0.001    -.0127245   -.0032961
       _cons |   1.568937   .4080033     3.85   0.000     .7692647    2.368608
-------------+----------------------------------------------------------------
    /lnalpha |   .4328395   .3201442                     -.1946316    1.060311
-------------+----------------------------------------------------------------
       alpha |   1.541629   .4935435                      .8231379    2.887268
------------------------------------------------------------------------------
Likelihood-ratio test of alpha=0:  chibar2(01) =   82.55 Prob>=chibar2 = 0.000
```

A3- Include only *manvn04* and exclude *pop04*

```
Negative binomial regression                    Number of obs   =         61
                                                LR chi2(5)      =      41.17
Dispersion     = mean                           Prob > chi2     =     0.0000
Log likelihood = -111.74867                     Pseudo R2       =     0.1555

------------------------------------------------------------------------------
  newmanfirm |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     manvn04 |   .0013518   .0012212     1.11   0.268    -.0010416    .0037453
   student04 |   3.03e-06   4.90e-06     0.62   0.536    -6.57e-06    .0000126
    gdpmil04 |  -6.02e-08   3.60e-08    -1.67   0.094    -1.31e-07    1.03e-08
        iz04 |   .3050211   .1091187     2.80   0.005     .0911523    .5188898
   harbordis |  -.0085718   .0024574    -3.49   0.000    -.0133881   -.0037554
       _cons |   1.667939   .4223932     3.95   0.000     .8400633    2.495814
-------------+----------------------------------------------------------------
    /lnalpha |    .575509    .310987                     -.0340144    1.185032
-------------+----------------------------------------------------------------
       alpha |   1.778035   .5529459                      .9665576    3.270793
------------------------------------------------------------------------------
Likelihood-ratio test of alpha=0:  chibar2(01) =   94.37 Prob>=chibar2 = 0.000
```

B. Conditional logit model

Similarly as in the negative binomial model, we have re-run some additional regressions inserting alternatively the variable *same industry* (*same4dgsic*) and the variable *Vietnamese firms* (*vnfirm4dgsic*) as well as between the variable *population* (*pop04*) and the variable *GDP* (*gdpmil04*). We always include in the regressions all control variables, except the variable *pop04* and find that the estimated results are robust and do not appear to result from collinearity amongst the regressors.

When we insert the variable related to the number of foreign firms in the same industry (*same4dgsic)*, the coefficient of the variable *vnfirm4dgsic*, numbers of Vietnamese firms in the same industry, becomes negative and statistically insignificant while there is no change for the other variables. This result shows that the positive correlation (0.60) between *same4dgsic* and *vnfirm4dgsic* is important. Vietnamese firms and foreign firms in the same industries tend to invest in the same locations. If we do not include the variable *same4dgsi* in the regression, its effect is attributed to *vnfirm4dgsic* giving a positive bias to the *vnfirm4dgsic* coefficient. Whenever we include *same4dgsic* variable, the coefficient of *vnfirm4dgsic* is negative and insignificantly different from zero. Moreover, by running the likelihood ratio tests we find that the models which omit the variable *same4dgsic* appear misspecified and are dominated by the models including it in the regressions.

It is also noted that the conclusions are the same if we always include in the regressions all control variables, except the variable *gdpmil04.*

B1 - Include only *vnfirm4dgsic* and exclude *pop04*

```
Conditional (fixed-effects) logistic regression   Number of obs   =      18216
                                                  LR chi2(5)      =    1384.81
                                                  Prob > chi2     =     0.0000
Log likelihood = -1237.6667                       Pseudo R2       =     0.3587

------------------------------------------------------------------------------
      choice |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
vnfirm4dgsic |   .0017871   .0004152     4.30   0.000     .0009734    .0026009
   student04 |   5.68e-06   4.07e-07    13.95   0.000     4.88e-06    6.48e-06
    gdpmil04 |  -1.70e-08   1.84e-09    -9.23   0.000    -2.06e-08   -1.34e-08
        iz04 |   .2958584   .0162552    18.20   0.000     .2639989     .327718
   harbordis |  -.0048122   .0012235    -3.93   0.000    -.0072102   -.0024142
------------------------------------------------------------------------------
```

B2- Include *vnfirm4dgsic*, *same4dgsic* and exclude *pop04*

```
Conditional (fixed-effects) logistic regression   Number of obs   =      18216
                                                  LR chi2(6)      =    1457.91
                                                  Prob > chi2     =     0.0000
Log likelihood = -1201.1171                       Pseudo R2       =     0.3777

------------------------------------------------------------------------------
      choice |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
vnfirm4dgsic |  -.0006223   .0005087    -1.22   0.221    -.0016193    .0003746
  same4dgsic |   .0262433   .0032533     8.07   0.000     .0198671    .0326196
   student04 |   6.09e-06   4.16e-07    14.64   0.000     5.28e-06    6.91e-06
    gdpmil04 |  -1.80e-08   1.88e-09    -9.59   0.000    -2.17e-08   -1.43e-08
        iz04 |   .2765016   .0167227    16.53   0.000     .2437256    .3092775
   harbordis |  -.0046295    .001217    -3.80   0.000    -.0070147   -.0022443
------------------------------------------------------------------------------
```

B3- Exclude *vnfirm4dgsic* and *pop04*

```
Conditional (fixed-effects) logistic regression   Number of obs   =      18216
                                                  LR chi2(8)      =    1522.52
                                                  Prob > chi2     =     0.0000
Log likelihood = -1168.8133                       Pseudo R2       =     0.3944

------------------------------------------------------------------------------
      choice |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
   forfirm04 |   .0021584   .0004492     4.80   0.000     .0012779    .0030389
  same4dgsic |   .0180987   .0025865     7.00   0.000     .0130294    .0231681
border4dgsic |  -.0081448   .0026788    -3.04   0.002    -.0133951   -.0028945
 samecountry |    .003184   .0008453     3.77   0.000     .0015274    .0048407
   student04 |   5.17e-06   4.45e-07    11.64   0.000     4.30e-06    6.05e-06
    gdpmil04 |  -2.37e-08   2.07e-09   -11.41   0.000    -2.77e-08   -1.96e-08
        iz04 |   .1747559   .0281758     6.20   0.000     .1195324    .2299794
   harbordis |  -.0039081   .0011754    -3.32   0.001    -.0062118   -.0016044
------------------------------------------------------------------------------
```

B4- Exclude *same4dgsic* and *pop04*

```
Conditional (fixed-effects) logistic regression   Number of obs   =      18216
                                                  LR chi2(8)      =    1479.23
                                                  Prob > chi2     =     0.0000
Log likelihood = -1190.4593                       Pseudo R2       =     0.3832

------------------------------------------------------------------------------
      choice |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
   forfirm04 |    .002459   .0004417     5.57   0.000     .0015933    .0033247
vnfirm4dgsic |   .0013815   .0004043     3.42   0.001     .0005891    .0021738
border4dgsic |  -.0089385    .002313    -3.86   0.000    -.0134719    -.004405
 samecountry |   .0036842   .0008468     4.35   0.000     .0020246    .0053438
   student04 |   4.77e-06   4.44e-07    10.74   0.000     3.90e-06    5.64e-06
    gdpmil04 |  -2.37e-08   2.07e-09   -11.47   0.000    -2.78e-08   -1.97e-08
        iz04 |   .1795017   .0282227     6.36   0.000     .1241862    .2348173
   harbordis |  -.0039263    .001169    -3.36   0.001    -.0062175   -.0016351
------------------------------------------------------------------------------
```

Notes:

We are also concerned about the fact that - a part from *iz04* and *harbordis* - all the regressors we include in the analysis are in absolute numbers, and they might all capture the effect of the size of the province. We thus have re-run regressions including the variables *iz04*; *harbordis*; *student* = student04/ pop04; *gdp_per capita* = gdpmil04/pop04; and with or without *gdpmil04*. The estimated results show that the original results of both the negative binomial model and conditional logit model are robust and the scale of GDP appears important rather than GDP per capita.

In sum, by re-running alternative regressions, we confirm that the estimated results presented in Tables 4 and 7 are robust and do not appear to result from collinearity among the regressors.

References

Baum, J. and Mezias, S. (1992), "Localized competition and organizational failure in the Manhattan hotel industry", *Administrative Science Quarterly*, vol. 37, pp. 580-604.

Baum, J. and Haveman, H. (1997), "Love thy neighbor? Differentiation and agglomeration in the Manhattan hotel industry, 1898-1990", *Administrative Science Quarterly*, vol. 42, pp.304-338.

Boudier-Bensebaa, F. (2005), "Agglomeration economies and location choice, foreign direct investment in Hungary", *Economics of Transition*, vol.13, no.4, pp. 605-628.

Cheng, L. and Kwan, Y. (2000), "What are the determinants of the location of foreign direct investment? The Chinese experience", *Journal of International Economics*, vol.51, pp.379-400.

Coughlin, C. and Segev, E. (2000), "Location determinants of new foreign-owned manufacturing plants", *Journal of Regional Science*, vol.40, pp.323-351.

Crozet, M., Mayer, T. and Mucchielli, J. (2004), "How do firms agglomerate? A study of FDI in France", *Regional Science and Urban Economics*, vol.34, pp.27-54.

David, P. and Rosenbloom, J. (1990), "Marshalian factor market externalities and the dynamics of industrial location", *Journal of Urban Economics*, vol.28, pp.349-370.

Dunning, J.H. (1993), "*Multinational enterprises and the global economy*", New York: Addison-Wesley.

Glaeser, E., Kalla, H., Scheinkman, J., and Schleifer, A. (1992), "Growth in cities", *Journal of Political Economy*, vol. 100, no.6.

General Statistics Office (2004), "*The situation of enterprises through the results of surveys conducted in 2001, 2002, 2003*", Hanoi: Statistical Publishing House.

General Statistics Office (2007), "*The situation of enterprises through the results of surveys conducted in 2004, 2005, 2006*", Hanoi: Statistical Publishing House.

General Statistics Office (2000-2005), "*Vietnam Statistical Yearbook*", CD-rom, Hanoi: Statistical Publishing House.

General Statistics Office website: http://www.gso.gov.vn/default.aspx?tabid=217

Guimaraes, P., Figueiredo, O. and Woodward, D. (2000), "Agglomeration and the location of foreign direct investment in Portugal", *Journal of Urban Economics*, vol. 47, pp.115-135.

Head, K., Ries, J. and Swenson, D. (1995), "Agglomeration benefits and location choice: evidence from Japanese manufacturing investments in the United States", *Journal of International Economics*, vol.38, pp.223-247.

Head, K., Ries, J. (1996), "Inter-city competition for foreign investment: static and dynamic effects of China'd incentive areas", *Journal of Urban Economics*, vol.40, pp.38-60.

Head, K., Ries, J. and Swenson, D. (1999), "Attracting foreign manufacturing: investment promotion and agglomeration", *Regional Science and Urban Economics*, vol.29, pp.197-218.

Kokko, A., Kotoglou, K., and Krohwinkel-Karlsson, A. (2003), "The implementation of FDI in Vietnam: an analysis of the characteristics of failed projects", *Transnational Corporations*, vol. 12, no. 3, pp. 41-77.

Krugman, P. (1991), "*Geography and Trade"*, Leuven: Leuven University Press and Cambridge (MA), London: the MIT Press.

Krugman, P. R. and Maurice, O. (1997), "*International Economics*", New York: Addison-Wesley.

Le Viet Anh (2004), "Locational determinants of foreign direct investment: the case of Vietnam", mimeo, Graduate School of International Development, Nagoya University.

Long, J.S. (1997), "*Regression models for categorical and limited dependent"*, Thousand Oaks, CA: Sage Press Publication

Long, J.S. and Freese, J. (2006), "*Regression models for categorical dependent variables using Stata"*, Texas: Stata Press Publication.

Marshall, A. (1920), "*Principles of Economics*", London: Macmillan.

McFadden, D. (1974), "Conditional logit analysis of qualitative choice behavior", in Zarembka P. (eds): *Frontiers in Econometrics*, New York: Academic Press.

Meyer, K. E. and Nguyen, H. V. (2005), "Foreign investment strategies and sub-national institutions in emerging markets: evidence from Vietnam", *Journal of Management Studies*, vol. 42, no.1.

Mirza, H. and Giroud, A. (2004), "Regional integration and benefits from foreign direct investment in ASEAN countries: the case of Vietnam", *Asian Development Economic Review*, vol. 21, no.1, pp.31-40.

MPI website: http://www.mpi.gov.vn/portal/page/portal/bkhdt/dtttnn(fdi)

Nguyen Ngoc Anh and Nguyen Thang (2007), "Foreign direct investment in Vietnam: an overview and analysis the determinants of spatial distribution across provinces", mimeo, Development and Policies Research Center.

Nguyen Thi Phuong Hoa (2002), "Contribution of foreign direct investment to poverty reduction: the case of Vietnam", source: http://www.zef.de/module/register/media/e086_FDI%20and%20poverty%20in%20Vietnam%20(Nov.%2002).pdf.

Nguyen Phi Lan (2006), "Foreign direct investment and its linkage to economic growth in Vietnam: a provincial level analysis", mimeo, Centre for Regulation and Market Analysis, University of South Australia.

Pham Hoang Mai (2002), "Regional economic development and foreign direct investment flows in Vietnam, 1988 – 1998", *Journal of the Asian Pacific Economy*, vol.7, no.2, pp.182-202.

Saxenian, A. (1994), "*Regional Advantage"*, Cambridge (MA): Harvard University Press.

Shaver, J.M. and Flyer, F. (2000), "Agglomeration economies, firm heterogeneity, and foreign direct investment in the United States", *Strategic Management Journal*, no.21, pp.1175-1193.

Wooldridge, J.M. (2003), "*Introductory Econometrics"*, Cincinnati, OH: South-Western College Publishing.

Yoffie, D. (1993), "Foreign direct investment in semiconductors", in Froot, K. (eds): *Foreign Direct Investment*, Chicago: University of Chicago Press.