Closing the Growth Gap: Entrepreneurship across Vietnam's Provincial Economies*

Chi L. Pham[1] and Stephan Weiler[2]

[1 2] Department of Economics, Colorado State University
Email: stephan.weiler@colostate.edu

Regional Economies and Policies (REP) 2016(1): 1 – 28.

Keywords *entrepreneurship, Vietnam, spatial econometrics, regional growth*

Abstract

This paper examines the effect of provincial characteristics on regional entrepreneurship growth in rapidly-evolving Vietnam. Combining theoretical endogenous growth models with spatially-explicit econometric techniques, a sequential series of regressions are run for the 63 provinces of Vietnam across the period of 2005 to 2013. The key findings are that a growth gap between the large-core-city and non-adjacent provinces persists, and the set of characteristics that have the greatest effect on provincial entrepreneurship growth are market-size oriented. Also, there is strong empirical evidence of spatial spillovers, implying that new firms and the development of a province's factors will have substantive impacts on entrepreneurial outcomes not only in the province itself but also in neighboring regions.

JEL Classification: L26, C30, R11

**This article has been developed from Chi L. Pham's unpublished graduate work at the Department of Economics, Colorado State University.*

Contents

1. INTRODUCTION

Vietnam's economy has grown rapidly over the last 20 years following the economic reform of the 1990s. The government's plan in 1986 to switch from the planned economy to the Socialist-oriented market economy, together with the Company Law enacted in 1990, and the Enterprise Law in 2002, have pushed Vietnam's entrepreneurship to a new level of development (Van Arkadie and Mallon, 2003). These pro-business conditions and laws have played an important role in helping Vietnam escape poverty, and promoting economic development.

Another stride making a drastic change in Vietnam's economy is joining the World Trade Organization (WTO) in 2007. The increase in exports has promoted the development of distribution and retail. Businesses have grown in both quality and quantity, with increasing kinds and standards of goods and services in multiple sectors. Nevertheless, most of the enterprises in Vietnam are still small.

Despite the remarkable contribution of entrepreneurship development to the economy, there seems to be an uneven growth between geographic regions of Vietnam. The two biggest provinces in Vietnam are Hanoi, the capital, and Ho Chi Minh City, which used to be the capital of the French colony of Cochinchina. It is highly possible that there exists a spillover effect from these two provinces to their neighbors, making the regions around them wealthier and while the rest is still growing slow (Hue, 2015). This fact has increased the income gap between the rich and the poor in Vietnam.

There is a tight relationship between entrepreneurship development and economic growth (Wennekers, 1999); more specifically, entrepreneurship development and income growth (Freeman, 1996; Oostendorp, 2009; Le, 2015). Accordingly, looking into how to promote enterprises' development in regions with low-income growth is extremely imperative for policy-makers to have a better guidance in increasing income in relatively poorer regions and reducing the income gap in Vietnam. The simple reality is that for most Vietnamese, income generation will require the entrepreneurial small enterprise development, as ubiquitous and widely accessible wage and salary jobs are not realistic for many years.

With these concerns, this paper attempts to find out which factors affect the growth of enterprises in a region and if there is truly a spillover effect. The authors begin with fixed effect, first difference of net change and percentage change to account for the biasedness and inefficiency problems of the pooled OLS regressions, setting a baseline for discerning the trend impact of those factors affecting entrepreneurship growth. Spatial Error, Spatial Lag, and Spatial Durbin models are then estimated to discover further the extent and effects of spatial correlation between neighboring regions.

The concluding Spatial Durbin Model provides perhaps the greatest number of controls for region-specific effects and multiple channels of spatial spillovers. For poorer regions of Vietnam, the number of firms in a province is positively and significantly affected by the number of firms in the neighboring provinces. On the independent variables effect, only variables that measure the market factor have economically significant effects on growth.

The next section presents the context and motivation including the proposed theoretical foundation of an endogenous growth model with entrepreneurship. The third section discusses the empirical modeling based on the noted theoretical foundations, including constituent variables, regression models, and data. The fourth section shows the model results and implications. The final part summarizes the key findings and offers policy conclusions.

2. CONTEXT

After two consecutive wars against France and the United States from 1858 to 1975, Vietnam was left devastated by poverty and hunger. Vietnam's new government built the economy towards centralized planning where resources were directly allocated; small businesses were eliminated, and the State controlled all of the economic activities. This mechanism proved ineffective as the Gross Domestic Product (GDP) growth rate continually dropped from 13.6% in 1977 to -3.5% in 1980 and inflation was going out of control (GSO).

Facing this crisis, the Government of Vietnam implemented an economic reform - Doi Moi (or 'economic renewal') in 1986, which showed effects in 1992 when the economy grew rapidly, and subdued inflation. Market economy replaced planning economy under a socialist orientation, where private businesses were allowed to operate, although the state sector remained the primary economic actor. After the implementation of Doi Moi in 1991 to 1995, average per capita GDP jumped to 6.6% and poverty was significantly reduced (Van Arkadie and Mallon, 2003).

The economy has developed considerably as Doi Moi policies significantly liberalized the market and mobilized resources for development that accelerated growth and controlled hyperinflation. Household businesses have played a significant role in rural and informal urban economic activities and have provided employment for most of the population. Formal private enterprises and foreign investors only became important economic actors as the transition progressed.

The 1992 Constitution set up an essential foundation for the private sector to compete with the state sector. There were 190 joint stock companies and 8,900 limited liability companies registered by 1996 (Van Arkadie and Mallon, 2003). In contrast, the number of state-owned enterprises declined in the period of 1989 to 2005 due to mergers, dissolutions, and acquisitions. The size of state-owned enterprises in GDP decreased in the period of 1994-2003, although it remained the largest sector (Meyer, 2006). After the creation of Enterprise Law in 2002, the number of newly registered private enterprises reached 36,000 in 2004 up from 14,457 in 2000. By June 2004, the total number of firms registered under the Enterprise Law reached 95,357 (Hakkala, 2007).

Another event which significantly contributed to the quick development of Vietnam's entrepreneurship is joining the World Trade Organization (WTO) in 2007. The percentage of GDP associated with export increased from 56.3% in 2005 to 65.3% in 2010 and reached the highest point in the past 40 years at 80.7% in 2014 (GSO). Implemented FDI increased 195.95% during 2006-2007. The percentage of the number of non-state enterprises increased from 96.17 % in 2005 to 99.14% in 2013.

2.1. Entrepreneurship and Economic Growth

Research on the relationship between economic growth and entrepreneurship growth has focused on the economic factors which affect business development and the effect of business development on the economy. Dejardin (2011) found a positive relationship between the development of new businesses and regional growth in Belgium. Bunten (2015) provided evidence for the significant effect of establishment births and deaths on employment growth across US counties.

Nonetheless, empirical studies on the economic determinants of new firm's formation have yielded diverse and even contradictory results. Guesnier (1994) and Armington (2002) found evidence of a positive effect of population change on new firm entry, while Audretsch (1994), Garofoli (1994) and Sutaria (2004) found none. On the other hand, while Audretsch (1994) and Wang (2006) found a positive impact of the change in the unemployment rate on new firm formation, Guesnier (1994), Garofoli (1994) and Sutaria (2004) found the impact to be negative. Finally, while Audretsch (1994) found no effect of the mean establishment size, Armington (2002) found a negative one and Sutaria (2004) found a positive one. These contradicting results make general foundations for policy development challenging.

Firms' location decisions alongside agglomeration and regional spillover effects are also interesting from a policy perspective. Looking at the determinants on the locations of firms in the United States, Coughlin (2000) found a positive effect of economic size, labor force quality, agglomeration economies, urbanization economies and transportation infrastructure on the location of new foreign-owned plants. Ellison (2007) stated that there was a significant and positive effect of transportation costs, labor pooling and technology spillovers on agglomeration in the US.

In Vietnam, the topic of agglomeration and determinants on the locations of firms is still new and has not been studied in depth, but would shed useful light on the drivers of enterprise development and their implications across regions. This paper in fact indirectly tests the importance of agglomeration economies by finding fundamental differences between large-core-cities and nonadjacent provinces, then using that distinction in its empirical analyses.

2.2. Entrepreneurship in Vietnam

On the issue of which firm characteristics increase firm growth, and survival length, Hansen (2009) found that in Vietnam, small firms grew faster than large firms, innovative firms survived longer, and firms that had government customers grew faster and survived longer. Significant evidence of growth originated from initial government support, tax exemption, and direct credit was also found. This suggests an important role of government support in entrepreneurship development in Vietnam.

Focusing on the relationship between the performance of incumbent firms and the net entry of new firms, Santarelli (2012) discovered that from 2000 to 2008, net entry of enterprises in Vietnam was associated with the performance of incumbent firms and the overall performance of the economy. His findings also suggested spatial spillovers between neighboring regions. Some research on small businesses does exist on selected Vietnamese regions. Freeman (1996) found a great jump in the number of small enterprises in Ho Chi Minh City – Vietnam's largest city after Doi Moi reform. He also suggested that small enterprises promoted both rural and urban economies by increasing income, providing cheap goods to the poor, and offering jobs to the lower class. Small enterprises helped staved off Vietnam's bankruptcy before the Doi Moi and helped the government discover an alternative path to national development. As entrepreneurship grew, nonfarm household enterprises in Vietnam had become important actors in the economy (Oostendorp, 2009). Agreeing with Freeman (1996), Oostendorp also found evidence that the nonfarm household enterprises development increased income, reduced inequality among households, and created jobs especially in rural areas.

Nevertheless, the role of the nonfarm household enterprise sector has been diminishing, particularly in urban areas. This trend was due to the liberalization after 1993 when the government made an effort to promote and facilitate the development of the private sector and focus on exports. A trade-off was made between developing the high-productivity and low-productivity sectors.

A divergence in the growth trends between the rich regions and the poorer regions in Vietnam is found using data of the number of enterprises in each province from 2005 to 2013 (GSO). The data is normalized against the province's population and separated into two groups: Group 1 includes Hanoi, Ho Chi Minh City, and their adjacent provinces; Group 2 includes the non-adjacent provinces. Figure I shows that the mean, min and max values of Group 1 are higher than those of Group 2, respectively. Moreover, through the years, Group 1 has grown faster than Group 2 in terms of the average number of firms per thousand people. This is a sign that the entrepreneurship growth gap between the rich and the poorer regions is increasing.

More detailed geographic data at the provincial level underscores the value of exploring subnational trends in small enterprise development. Figure II presents a map of the average number of enterprises across provinces in 2008. Enterprises in Vietnam mainly agglomerate in Hanoi, Ho Chi Minh City, their neighboring provinces, and coastal provinces. The remaining provinces have sparsely located enterprises, especially in the North-West mountain region and some provinces south of Ho Chi Minh City where floods frequently occur.

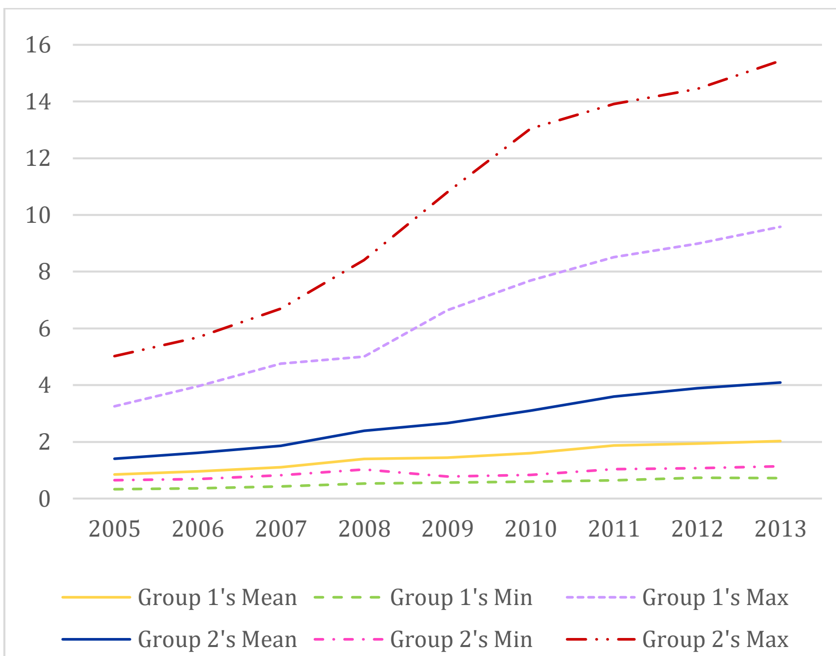

Figure I. Average Number of Enterprises per Thousand People 2005-2013

Source: Data from Statistical Year Book of Vietnam (GSO).

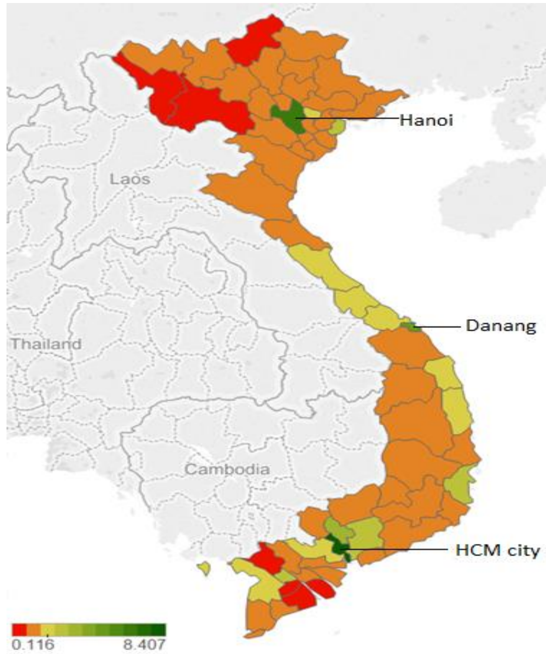

Figure II. Average Number of Enterprises per Thousand People 2008

Source: Data from Statistical Year Book of Vietnam 2009 (GSO); HCM: Ho Chi Minh.

Overall, research on regional entrepreneurship growth and spillover effects between regions in Vietnam is still limited. Due to data limitations, most research only focuses on case studies. This paper will look at entrepreneurship growth in Vietnam as a whole - including all sizes of enterprises - and also account for possible spatial spillovers.

2.3. Endogenous Growth Model with Entrepreneurship

Vietnam's regional entrepreneurial growth can be framed with a regional endogenous growth model, as other growth-oriented papers have done (e.g. Bunten et al., 2015). One of traditional Solow-type exogenous growth models' key implications is that regional convergence should occur over the longer-run. But it has been proven to be false in some cases (Romer, 1994). Noticing the flaws of the traditional neoclassical Solow growth theory, Romer (1994) built a model where he dropped two assumptions of the exogenous growth model. The first assumption states that technological change is exogenous while the second one purports same technological opportunities are available in all countries. The behavior of an economy then becomes endogenous and can be represented by the following equation:

$$(1)\quad \hat{y} = \alpha\hat{k} + \hat{A}$$
$$= \alpha\left[sA^{1/\alpha}y^{(\alpha-1)/\alpha} - n\right] + \hat{A},$$

where "^" denotes the exponential growth rate of a variable.

Since this is a closed economy by assumption, the saving rate s will also be equal to the investment rate. The second line of equation (1) shows that outside of the steady state, a change in the investment rate and the level of output per labor will change the growth rate. This is synchronous with the finding of Martin (1998) that investment in physical capital is strongly correlated with, and causally related to, growth.

Applying the endogenous growth model above to a provincial scale, a model is built based on Romer's endogenous growth model (Romer, 1994) and Mankiw's model (Mankiw, 1992) where A is not the same across provinces but is determined locally. Human capital, which accounts for the skills, knowledge, and experience of the population, needs to be incorporated in the growth model as it is less mobile than physical capital and therefore is a key component of the potential and competitiveness of a province. Moreover, the growth of enterprises in a province not only is affected by enterprises' inputs (physical capital, labor and human capital) but also depends on exogenous provincial characteristics. Infrastructure and facilities development of a province could also affect the development of firms. For example, an improvement in the transportation system or auxiliary facilities (warehouses, airports, hospitals) in a province will attract firms to locate in that province. On the other hand, market factors are undeniably important to the location and development of firms. A province with a big population implies a big market for firms. Population income of a province measures for the purchasing power of the market. Government policy can also be an important factor accounting for business growth in Vietnam (Hansen, 2009), whether indirectly via the above channels or through direct investments of its own.

Regarding the model, assume that an increase in investment in capital will not only increase the physical capital but also increase the level technology through knowledge spillovers. Secondly, assume that an increase in the labor amount will have a negative spillover effect on the labor-saving innovations. Thirdly, assume an increase in the education level of the population will increase the technological advancement in a province. Incorporating exogenous and endogenous elements, the output function for a province would be:

$$(2)\ Y_i = A(K,H,L,F,M,P)K^{\alpha}H^{\beta}L^{(1-\alpha-\beta)} \qquad \alpha,\beta \in (0,1);\ i \in [1,63],$$

where H is human capital, F is provincial infrastructure and facilities; M is market and P is government policy. The i indicates the individual province. All the variables excluding Y are both of the individual province and of every other province to account for the effect from the individual province and its neighbors. The element A in function (2) indicates Total Factor Productivity (TFP) which is the portion of output not explained by the amount of inputs used in production, i.e. the residual after accounting for all physical capital, human capital, and labor inputs. It is a function of endogenous factors K, H, L and exogenous factors F, M, P. For simplicity, suppose $\mathrm{A}(K,H,L,F,P,M) = K^{x}H^{y}L^{z}FMP$ with $x, y, z > 0$. Plugging this into function (2) to make A disappear, the results will be:

$$(3)\ \mathrm{Y}_i = K^{(\alpha+x)}H^{(\beta+y)}L^{[1-\alpha-\beta+z]}FMP.$$

In this case, growth is only a function of K, H, L, F, M, and P.

The factors K, H, L, F, M and P are expected to have effects on growth Y. An increase in K, H, L, and M would increase growth Y. An increase in the number of provincial facilities F that are used by firms or support firms' activities would increase firms' growth. An increase in pro-enterprise development policy P would increase growth Y. Any other investments in provincial facilities and policy that are not pro-enterprise development would cause growth to decrease.

Most studies on regional growth use pooled data for all geographical areas in the system studied, which assumes that the convergence process is identical across all regions. This is not usually true as the rate of convergence varies across regions, i.e. different regions may converge to different growth level that reflects local differences in structural characteristics (Martin, 1998). This paper tries to account for the problem by incorporating fixed effects into the model, both explicitly as well as implicitly via first-differences.

The second key issue which must be incorporated into statistical modeling is that the economics of regions might be interrelated where the growth of a region might depend on the growth of other regions. Moreover, clusters of high- and low-growth regions might emerge (Martin, 1998). The authors use a variety of spatial regression techniques to address these interrelationships and clustering effects.

3. EMPIRICAL MODELING OF VIETNAM'S ENTREPRENEURSHIP

3.1. Variables

Ideally, the growth of entrepreneurship in a province would be quantitatively evaluated across all firms as well as growth among existing firms themselves. Nonetheless, it is difficult to obtain data on the size change of each enterprise in Vietnam. For this reason, the number of enterprises in a province will be its key measure of entrepreneurship growth and is chosen to be the dependent variable.

Independent variables are chosen based on theory and the availability of data. Variables which account for physical capital K are material inputs of the food industry (specifically, production of cereal, aquaculture, livestock, and poultry), Foreign Direct Investment (FDI), and the revenue of enterprises (which will be counted as the capital of the next production cycle). Variables which account for human capital H are the normalized concentration of college students, high school students (from 10th grade to 12th grade) and secondary and primary students (from 1st grade to 9th grade). Nevertheless, these numbers do not measure the human capital factor accurately as many students from poor regions move to major cities to attend better high schools and colleges. Furthermore, these data are simply enrolled students, not the educational level of actual business owners and workers, which are the most relevant actors. Data on the percentages of the workforce with specific skill/educational qualifications, or at least average educational attainment, would be ideal, but are not yet available at the province level.

The variable which accounts for labor L is the number of employees working in enterprises. The variable which proxies for infrastructure and facilities F is the number of hospitals. In terms of market factors, population and net migration are measures of market size and change while retail sales of goods and services is the measure of market purchasing power. The final variable, volume of freight transported, accounts for both physical capital K (volume of input materials transported to firms) and market M (volume of outputs transported to the market). No variable which suitably accounts for government pro-entrepreneurship development policies that are different across provinces can be found. Hence, policy effects will be included in the error term of the model. Table I identifies the information of the explanatory variables and their expected signs based on theory.

Table I. Explanatory Variables and Expected Signs

Explanatory Variables	*Expected Signs*	*Factors*
Production of Material Inputs of Food Industry	+	*K*
Foreign Direct Investment (FDI)	+	*K*
Revenue of Enterprises	+	*K*
Number of College Students	+	*H*
Number of High School Students	+	*H*
Number of Secondary and Primary Students	-	*H*
Number of Employees in Enterprises	+	*L*
Population	+	*M*
Net Migration	+	*M*
Retail Sales of Goods and Services	+	*M*
Number of Hospitals	+	*F*
Volume of Freight Transported	+	*K/M*

3.2. Econometric Models

3.2.1. Fixed Effect and First Differences

Given the noted likelihood of substantial time-invariant regional differences, pooled analyses are likely to have significant omitted variable biases. To fix the bias problem of pooled OLS, fixed effect regressions are estimated allowing the individual province to have its intercept. This helps the authors study the theory that different regions may converge to different growth level.

Another type of regression which also accounts for pooled OLS's bias and inefficiency is the first difference regression. Moreover, it also takes care of the omitted variable bias assuming that the omitted variable is unchanged through time. First difference regressions with net change allow a focus on the flows. First difference regression with percentage change normalizes the trend of the changes in the independent variables and their effect on the trend of growth.

Therefore, by the reduced form growth model (3) above, the empirical strategy for static Fixed Effect and dynamic First-Difference OLS regressions are expressed by equation (4) and (5), respectively.

(4)

$$\begin{aligned} Growth_{it} = \alpha_i &+ \beta_1 Pop_{it} + \beta_2 NetMigration_{it} + \beta_3 FDI_{it} + \beta_4 RetailSale_{it} \\ &+ \beta_5 FreightVolume_{it} + \beta_6 Labor_{it} + \beta_7 Revenue_{it} \\ &+ \beta_8 CollegeStudents_{it} + \beta_9 HighschoolStudents_{it} \\ &+ \beta_{10} PrimarySecondaryStudents_{it} + \beta_{11} Hospital_{it} + \beta_{12} Cereal_{it} \\ &+ \beta_{13} Aquaculture_{it} + \beta_{14} Livestock_{it} + e_{it} \end{aligned}$$

$$i \in [1,63];\ t \in [2008,2013]$$

Where α_i is the unobserved time-invariant individual effect, *Growth* is the number of acting enterprises, *Pop* is the province's population, *NetMigration* is the difference between the percentage of population moving in and the percentage of population moving out of a province, *FDI* is the total foreign direct investment registered, *RetailSale* is the retail sales of goods and services at current prices, *FreightVolume* is the volume of freight transported, *Labor* is the number of employees working in enterprises, *Revenue* is the net revenue of enterprises, *CollegeStudents* is the number of students in college, *HighschoolStudents* is the number of students in high school, *PrimarySecondaryStudents* is the total number of students in primary school and secondary school, *Hospital* is the number of hospital, *Cereal* is the production of cereal, *Aquaculture* is the production of aquaculture, *Livestock* is the total headcounts of livestock and poultry raised, and e_{it} is the error term.

(5)

$$\begin{aligned} \Delta Growth_{it} &= Growth_{it} - Growth_{it-1} \\ &= \beta'_1 \Delta Pop_{it} + \beta'_2 \Delta NetMigration_{it} + \beta'_3 \Delta FDI_{it} + \beta'_4 \Delta RetailSale_{it} \\ &\quad + \beta'_5 \Delta FreightVolume_{it} + \beta'_6 \Delta Labor_{it} + \beta'_7 \Delta Revenue_{it} \\ &\quad + \beta'_8 \Delta CollegeStudents_{it} + \beta'_9 \Delta HighschoolStudents_{it} \\ &\quad + \beta'_{10} \Delta PrimarySecondaryStudents_{it} + \beta'_{11} \Delta Hospital_{it} + \beta'_{12} \Delta Cereal_{it} \\ &\quad + \beta'_{13} \Delta Aquaculture_{it} + \beta'_{14} \Delta Livestock_{it} + \Delta e_{it} \end{aligned}$$

$$i \in [1,63];\ t \in [2008,2013]$$

3.2.2. Spatial Error, Spatial Lag, and Spatial Durbin Models

As noted in the literature review and the theory sections, there is a great chance of spillover and clustering effect among provinces. When variables for each province are mapped, the dependent variable, as well as most of the independent variables, clearly cluster around two biggest cities, Hanoi in the north and Ho Chi Minh City in the south. Hence, the authors apply spatial econometric techniques to account for the effect of the growth of enterprises in neighboring provinces on the growth of enterprises in a province.

Taking a look at the Spatial Error Model (SEM) with fixed effects is necessary as it corrects the potentially biased influence of spatial autocorrelation due to possible missing of important variables. There are reasons to believe that the regression has missing variables with distinct spatial footprints (regional government policies, energy resources, transportation facilities, supporting industries and services, etc.). Hence, running a fixed effect or first different regression alone will yield a biased result. SEM fixes this problem by spatially lagging the error term to account for missing variables' spatial autocorrelation.

In Vietnam, the number of firms in a province is actively influenced by its neighbors because of the regional spillover effects. The Moran's I test for the dependent variable (Table V below) further addresses this belief and shows evidence of spatial autocorrelation between the growth of each province with a 1% significance. This proposes the usefulness of a Spatial Lag Model (SLM). SLM is another form of regression that addresses the spatial autocorrelation problem by running a spatial lagged dependent regression. Instead of looking at the error term like SEM, SLM looks at the spatial autocorrelation of the dependent variable.

Since both SEM and SLM are believed to be useful, Spatial Durbin Model, which incorporates the spatial interaction between both the dependent and independent variables, is chosen as the most suitable model to apply in the regional model of Vietnam. Base on theory and the Moran's I tests (Table V and related discussion below), it is likely that the spatial spillover effects are not only in the dependent variables but also the independent variables (market size and purchasing power, capital and labor spillovers). This model looks at the simultaneous feedback effect between provinces where a change in the explanatory variable of province i affects the dependent variable of the neighbor provinces, the neighbors of the neighboring provinces, and so on.

Furthermore, as further shown, Chow Test results indicate that the two datasets generate statistically different coefficients, indicating that they react differently to core provincial characteristics. Therefore, the authors divide the dataset into two sets; the first includes Hanoi and Ho Chi Minh City's non-adjacent neighbors (47 provinces, n_1=423); the second includes Hanoi, Ho Chi Minh City, and their adjacent neighbors (16 provinces, n_2=144).

3.3. Data

The geographic scale of this analysis is at the province-level, as more spatially detailed data are not available. Enterprises and provincial yearly data are taken from Vietnam General Statistical Office (GSO). These are the panel-type data which span from 2005 to 2013 across 63 provinces. There might be some measurement errors due to the limitations of transparency and advanced measurement methodology of the Statistical Office of Vietnam. Unfortunately, there is no alternative data source that the authors can use to compare the quality of data. Because the topic is relatively new in Vietnam, little studies were found that used similar dataset and no study could be used as a comparison. Since provinces in Vietnam vary greatly in size, all the variables are normalized. The dependent variable is normalized against provincial area (square kilometer), which the Hausman specification test suggests to being a better normalization than population (see Appendix 4 for test and Appendix 5 for population-normalized results). The independent variables are normalized against the provincial population, except the volume of freight transported, the number of employees in enterprises and business revenue which are normalized by the number of enterprises for better interpretation. From 2008 onwards, provinces of Hanoi and Hatay are merged. Therefore, variables before 2008 are sums of two provinces and are looked at as variables of Hanoi.

Table II. Dependent Variable, Independent Variables and Their Units

	Code	Unit	Factor
Dependent Variable			
Number of Enterprises per Thousand People	*Growth*	*Enterprises/Square Kilometer*	
Independent Variables			
Population	*Pop*	*People*	*M*
Net Migration	*NetMigra*	*Percent*	*M*
FDI per capita (P)	*FDI*	*Thousand USD/Person*	*K*
Retail Sales per capita (P)	*Retailsale*	*Million VND/Person*	*M*
Volume of Freight Transported (B)	*Freight*	*Thousand Tons·Kilometer/Enterprise*	*K/M*
Average Labor in an Enterprise (B)	*Labor*	*People*	*L*
Average Revenue of an Enterprise (B)	*Revenue*	*Billion VND/Enterprise*	*K*
Ratio of College Students to Population (P)	*College*	*(Ratio)*	*H*
Ratio of High School Students to Population (P)	*High*	*(Ratio)*	*H*
Ratio of Second. Primary Students in Population (P)	*Primsec*	*(Ratio)*	*H*
Number of Hospital per Thousand People (P)	*Hospital*	*Hospital/Thousand People*	*F*
Cereal Production per capita (P)	*Cereal*	*Kilograms/Person*	*K*
Aquaculture Production per capita (P)	*Aqua*	*Kilograms/Person*	*K*
Population of Livestock and Poultry per capita (P)	*Livestock*	*Heads/Person*	*K*

Note: *(P)* Normalized by population; *(B)* Normalized by number of business.

Summary statistics in Table III underscore the differences between the Large-Core-City and Non-Adjacent provinces, the mean of the average number of enterprises per square kilometer in Hanoi-Ho Chi Minh City-and-neighbor regions is 4.176 while that in the non-adjacent provinces is only 0.557. Overall, there are signs that the entrepreneurship in Hanoi, Ho Chi Minh City and their neighboring provinces is more key developed than the entrepreneurship in non-adjacent provinces. It is also likely that the types of enterprises in these major cities are different from those elsewhere in Vietnam, but the data do not allow exploration of such distinctions. This line of inquiry can be left to future research.

Table III. Summary Statistics

Non-Adjacent Provinces (N = 423)					HN, HCM and Neighbors (N = 144)				
Variable	Mean	Std. Dev.	Min.	Max.	Variable	Mean	Std. Dev.	Min.	Max.
Growth	0.557	0.951	0.021	7.356	Growth	4.176	9.928	0.119	57.608
Pop	1159.834	613.382	288.4	3477.7	Pop	1971.055	1874.078	778	7820
NetMigra	-2.687	5.394	-27.300	36.200	NetMigra	3.105	12.976	-11.800	74.600
FDI	0.184	0.993	0.000	17.479	FDI	0.444	1.053	0.000	9.532
Retailsale	11.708	8.482	1.119	49.319	Retailsale	15.853	14.413	1.750	79.161
Freight	0.403	0.425	0.021	2.683	Freight	0.288	0.139	0.043	0.651
Labor	36.088	14.706	14.110	108.276	Labor	52.746	24.462	19.884	150.333
Revenue	17.140	11.325	2.603	85.277	Revenue	34.777	29.268	3.804	186.265
College	0.019	0.058	0.000	0.498	College	0.050	0.112	0.001	0.539
High	0.034	0.009	0.010	0.058	High	0.034	0.007	0.013	0.053
Primsec	0.170	0.033	0.112	0.480	Primsec	0.144	0.022	0.090	0.200
Hospital	0.013	0.005	0.007	0.033	Hospital	0.011	0.003	0.005	0.018
Cereal	572.770	447.970	34.800	2578.800	Cereal	446.900	363.298	11.700	1931.100
Aqua	100.752	125.356	0.537	588.779	Aqua	48.871	67.314	4.805	288.230
Livestock	3121.281	1295.854	343.481	7877.005	Livestock	4347.670	2364.395	26.378	9922.283

Note: All variables are normalized. HN, Hanoi; HCM, Ho Chi Minh City.

4. EMPIRICAL RESULTS

4.1. Diagnostic Tests

Evidence of heteroskedasticity is found after running the Breusch-Pagan / Cook-Weisberg test for the following null hypothesis:

H_0: Constant variance.
$\chi^2(1) = 519.16$
$p = 0.0000$.

Therefore, H_0 is rejected. All regressions are tested for and have heteroscedasticity; hence, all the regressions below are run with robust standard errors to account for heteroskedasticity.

The correlation matrix of all the independent variables shows no serious problem of collinearity (Appendix 1). Variance inflation factors also show little sign of multicollinearity (Appendix 2).

A Chow test is run, and the result suggests that the coefficients in the linear regressions on two datasets: the non-adjacent provinces and Hanoi, Ho Chi Minh City with their adjacent provinces, are significantly different, which suggests that the effects of the growth factors are not the same across regions. This test further indicates that splitting the datasets into two smaller subsets gives more robust results than putting all provinces together. Splitting the datasets makes sense empirically since Hanoi, Ho Chi Minh City and their adjacent provinces are significantly different from the non-adjacent provinces when plotting the dependent and independent variables. The growth path of the large-core cities is expected to be remarkably different from that of non-adjacent provinces.

Table IV. Chow Test Results

Model	***F***	***df***	***p***
Pooled OLS	*17.5205*	*15/537*	*0.00*
Fixed Effect	*15.3308*	*15/537*	*0.00*
First Difference	*11.1397*	*15/474*	*0.00*

An F-test for fixed effects shows that there are fixed effects among provinces in Vietnam. It is thus appropriate to run a fixed effect model instead of the pooled OLS model for the following null hypothesis:

H_0: the individual intercepts of each province are all zero, i.e. $\alpha_i = 0$ ($i \in [1,61]$).
$F\,(60, 474) = 40.41$.
$p = 0.0000$.

As a result, H_0 is rejected.

Moran's I tests for spatially lagged dependent and independent variables show that there is spatial autocorrelation between dependent and independent variables among regions with the confidence level of 5% for the two-tailed test of the following null hypothesis. Accordingly, the Spatial Durbin Model, which looks at spatial autocorrelation of both dependent and independent variables, is likely to be the most appropriate ultimate model.

H_0: the coefficient on the spatially lagged variable is zero.

Table V. Moran's I Test Results

Variable	***Moran's I***	***Z***	***p (2 tailed)***
Growth	*0.148*	*14.875*	*0.000*
Pop	*0.263*	*26.147*	*0.000*
NetMigra	*0.067*	*6.939*	*0.000*
FDI	*0.015*	*2.015*	*0.044*
Retailsale	*0.175*	*17.413*	*0.000*
Freight	*0.34*	*33.927*	*0.000*
Labor	*0.371*	*36.936*	*0.000*
Revenue	*0.173*	*17.495*	*0.000*
College	*-0.023*	*-2.155*	*0.031*
High	*0.534*	*52.748*	*0.000*
Primsec	*0.304*	*30.495*	*0.000*
Hospital	*0.511*	*50.548*	*0.000*
Cereal	*0.631*	*62.503*	*0.000*
Aqua	*0.689*	*68.172*	*0.000*
Livestock	*0.417*	*41.222*	*0.000*

4.2. Regression Results

4.2.1. Fixed Effects (Static) and First Differences (Dynamic)

Variables which account for the growth factor *M* are significant overall, with the estimators for population and retail sales per capita being positive and mostly significant for both datasets in all regressions. The magnitude of the effect varies across regions with the strongest effect being in the Large-Core-City provinces. The coefficient of net migration is negative and statistically significant for the second dataset for fixed effect and percentage change regressions. This means that an increase in the net change of the net migration of a province will lead to fewer firms located in that province. It is possible that in crowded provinces like Hanoi and Ho Chi Minh City, an increase in the congestion could decrease the available area for firms to locate.

Table VI. Fixed Effect and First Difference Regression Results

	Fixed Effect		First Difference (Net Change)		First Difference (Percentage Change)	
	Nonadjacent	Large-Core-City	Nonadjacent	Large-Core-City	Nonadjacent	Large-Core-City
Intercept	*-1.6356	***-116.74	*-0.0440	**-0.4004	**0.0285	***0.0894
Pop	**0.0028	***0.0181	**0.0035	***0.0152	0.5494	0.5467
NetMigra	0.005	**-0.0726	0.0011	-0.0032	0.0001	**-0.0022
FDI	0.0099	0.0248	0.0009	0.0156	0.0001	**-0.0007
Retailsale	***0.0499	***0.1685	***0.0331	**0.1903	***0.3595	**0.1937
Freight	-0.0796	-1.6564	***-0.3425	**-2.4664	***-0.1541	***-0.4059
Labor	0.0054	***0.0731	**-0.0045	0.0232	***-0.6364	***-0.4673
Revenue	***-0.0084	-0.0097	-0.0028	-0.0008	***-0.0857	-0.0255
College	***2.3640	0.7822	0.4367	-0.5796	***-0.0022	*0.0059
High	-0.6856	**97.4415	0.5166	*70.4093	*0.1051	***0.4269
Primsec	0.9615	**26.4778	-0.2781	4.8226	-0.0622	*-0.2777
Hospital	-7.8611	***547.073	1.52	109.283	0.0039	-0.0659
Cereal	***-0.0013	-0.0004	**-0.0002	0.0032	-0.029	-0.0558
Aqua	-0.0003	-0.011	0.0008**	0.0061	0.0218	0.0601
Livestock	-0.0001	-0.0001	-0.0001	0.0001	0.0299	0.0538
N	423	144	423	144	423	144
Provinces	47	16	47	16	47	16
F-value	***47	***88.9	***3.19	***5.01	***19.3	***11.89

Note: The dependent variable is the average number of enterprises per square kilometer, 2005-2013. Estimates with one, two, three stars are statistically significant at the 10 percent, 5 percent, and 1 percent levels, respectively, using a two-tailed test. HN, Hanoi; HCM, Ho Chi Minh City.

With regards to the human capital factor H, the coefficient on the percentage of college students among population is statistically significant at 1% level for non-adjacent provinces in the fixed effect regression where it is positive, and in the percentage change regression where it is negative. The overall trend is that in recent years, the unemployment rate has decreased in all regions in Vietnam; however, that rate in mountain and central areas is decreasing faster than that in the plain areas around Hanoi and Ho Chi Minh (GSO). In other words, there is a higher chance of graduating college students to find an entrepreneurial niche in non-adjacent provinces than in regions around Hanoi and Ho Chi Minh City. Nevertheless, that effect seems to be slowing down since the estimator is negative in the percentage change regression. The relationship between the percentage of students in high school and the number of businesses is statistically significant only for the second dataset in the fixed effect and the net change regression where it is positive. It is positive and statistically significant for both datasets in the percentage change regression, indicating that increased concentrations of high school students are leading to more enterprises in those provinces. The coefficient of the percentage of primary and secondary students is statistically significant for the second dataset in the fixed effect regression where it is positive, and in the percentage change regression where it is negative. The positive sign is not the expected sign according to the hypotheses.

On the other hand, the measure for the labor factor L is positive and statistically significant at 1% level for the fixed effect regression of the second dataset but is negative and statistically significant for both first difference regressions. Larger firms exist in areas with more firms, but areas with increasing average size of firms reduce the concentration of firms, suggesting that growing firms absorb smaller firms' markets.

The physical capital factor K does not show signs of a significant effect on firms' growth in all regressions. The parameter estimate of FDI per capita is only statistically significant for the second dataset in the percentage change regression. The volume of freight is negative and statistically significant in both first difference regressions. The estimator for enterprises' average revenue is negative for the non-adjacent provinces in the fixed effect and percentage change regressions while in other datasets, it is not significant. Surprisingly, the estimators for the average production of cereal and aquaculture per capita is negative and statistically significant for the non-adjacent regions, suggesting that large amounts of primary food production may substitute for and crowd out enterprise development. In contrast, the measure for the facilities factor F is positive and statistically significant for the second dataset in the fixed effect regression, implying that a hospital in a province makes the area more attractive to enterprise development.

4.2.2. Spatial Error Model

The SEM in Table VII accounts for possible missing variables which are spatially correlated using spatially lagged error terms.

Moving to the models that properly incorporate inter-provincial spatial effects, the authors also add a benchmark for economic significance alongside the usual focus on statistical significance. Given the standard deviation of the average number of enterprises per square kilometer in Table III, a reasonable target for "economically significant" (versus simple statistical significance) provincial entrepreneurship growth in non-adjacent provinces is an increase one firm per square kilometer per year. The interpretation, henceforth, will pay particular attention to the non-adjacent provinces for better interpretation of which factors affect the entrepreneurship growth in poorer regions of Vietnam.

The regression results in Table VII further highlight the significant effect of the factor M on provincial growth. The population has a positive and economically significant coefficient in both datasets. NetMigra and Retailsale have positive and economically significant coefficients for the first dataset. The estimator for the F factor is positive and economically significant for the second dataset which suggests the number of hospitals per thousand people influences entrepreneurship growth of Large-Core-City provinces. The factor L also has a positive and economically significant coefficient. The H and K factors show little sign of significant influence on regional growth as coefficients of College, High, Primsec, FDI, Aqua, and Livestock are not statistically significant. More importantly, there is a significant effect of spatial error terms as lambda is positive and statistically significant for the first dataset. This suggests that there are missing variables which are spatially correlated for the non-adjacent provinces.

Table VII. Spatial Error Model Results

	Nonadjacent	*Large-Core City*
Pop	**0.0047*	****0.0182*
NetMigra	**0.0058*	*-0.0623*
FDI	*-0.0025*	*0.0486*
Retailsale	****0.0558*	*0.1567*
Freight	*-0.3550*	*-2.1960*
Labor	*0.0091*	***0.0750*
Revenue	*-0.0084*	*-0.0111*
College	*1.1134*	*0.8129*
High	*-9.2192*	*83.8756*
Primsec	*1.6381*	*28.7847*
Hospital	*-7.1488*	****558.628*
Cereal	****-0.0010*	*-0.0009*
Aqua	*-0.0009*	*-0.0046*
Livestock	*-0.0001*	*-0.0001*
λ	****0.5232*	*-0.1362*
σ^2_e	****0.0457*	****1.7951*

Note: The dependent variable is the average number of enterprises per square kilometer, 2005-2013. Estimates with one, two, three stars are statistically significant at the 10 percent, 5 percent, and 1 percent levels, respectively, using a two-tailed test. λ is the coefficient of the spatially lagged error terms. HN, Hanoi; HCM, Ho Chi Minh City.

4.2.3. Spatial Durbin Model

After finding evidence of spatial autocorrelation through the Moran's I test and spatially mapping all the variables, a Spatial Durbin Model which incorporates fixed effects with spatially lagged dependent and independent variables was estimated. The main results are reported in Table VIII with the average number of enterprises per square kilometer being the dependent variable. Another result of the SDM with the dependent variable normalized by provincial population is shown in Appendix 5. Both normalization methods, by population and by area, are reasonable since naturally, there would be more firms in populated or large provinces. Based on the Hausman specification test (Appendix 4), the SDM regression with area normalization is a better fit since normalizing against provinces' population yields inconsistent results while normalizing against provinces' area does not. For this reason, the interpretation of the SDM will focus on the area normalization method.

Table VIII. Spatial Durbin Model Results with Area Normalization in the Dependent

	Total *(= Direct + Indirect)*		***Direct Effect***		***Indirect Effect***	
	Nonadjacent	***Large-Core***	***Nonadjacent***	***Large-Core***	***Nonadjacent***	***Large-Core***
Pop	*-0.0044	0.0044	*0.0038	***0.0125	***-0.0082	***-0.0081
NetMigra	0.0145	**0.0582	***0.0078	-0.0200	0.0067	***0.0782
FDI	0.0307	0.1920	0.0055	-0.0574	0.0252	0.2495
Retailsale	**0.0562	***0.2369	***0.0758	***0.4102	*-0.0196	***-0.1733
Freight	***0.6947	**-5.2895	*-0.3448	*-2.4233	***1.0396	-2.8661
Labor	-0.0121	0.0069	0.0054	***0.0495	-0.0175	-0.0425
Revenue	-0.0115	-0.0067	-0.0077	**-0.0125	-0.0037	0.0057
College	*2.9819	-1.1329	1.2236	**-3.2122	*1.7583	2.0793
High	5.7835	*82.9589	*-8.5762	57.8507	14.3597	25.1081
Primsec	1.5226	***59.6270	1.4430	10.0156	0.0795	*49.6114
Hospital	63.1306	**450.507	**24.2457	66.8557	38.8849	**383.651
Cereal	**-0.0017	*-0.0058	**-0.0006	*0.0047	-0.0011	***-0.0105
Aqua	0.0019	***0.1479	0.0000	***0.0995	*0.0018	0.0484
Livestock	0.0000	**0.0004	-0.0000	0.0000	0.0000	*0.0004
ρ	***0.3443	***-0.3791	***0.3443	***-0.3791	***0.3443	***-0.3791
σ^2_e	***0.035	***0.6675	***0.0353	***0.6675	***0.0353	***0.6675
N	549	549	549	549	549	549
Provinces N	61	61	61	61	61	61
Mean FE	0.6916	0.9560	0.6916	0.9560	0.6916	0.9560
R^2 Within	0.0159	0.8183	0.0159	0.8183	0.0159	0.8183
R^2 Between	0.0066	0.7501	0.0066	0.7501	0.0066	0.7501
R^2 Overall	3.0522	3.4090	3.0522	3.4090	3.0522	3.4090

Note: The dependent variable is the average number of enterprises per square kilometer, 2005-2013. Coefficient with one, two, three stars are significant at the 10 percent, 5 percent, and 1 percent levels, respectively, using a two-tailed test. “rho” is the coefficient of the spatially lagged dependent variable. The Direct effect is the effect of the original province’s Independent variables on that province’s growth taking feedback effect into account. The Indirect effect is the effect of the independent variables of a province’s neighbors on that province’s growth taking feedback effect into account. HN, Hanoi; HCM, Ho Chi Minh City.

The coefficient “rho” is significant for both datasets, which leads to the belief that the number of firms in a province is affected by the number of firms in its neighboring provinces. Taking the feedback effect into account, there is evidence that the strongest factor affecting entrepreneurship growth in a province is, again, the M factor. For all regions in Vietnam, entrepreneurship growth in a province is positively affected by the population of that province and negatively affected by that of the neighboring province. An increase of 263 people in a non-adjacent province is related to a one firm increase per square kilometer. On the other hand, an increase of 122 people in that province’s neighboring province is related to a decrease of one firm per square kilometer, ceteris paribus. Taking the standard deviation of the variable Pop from Table III (which is 613.382 people) as a reference, the direct and indirect effects of the population in non-adjacent regions are both economically significant. The total effect, however, has a negative and statistically significant coefficient for the non-adjacent provinces, which suggests that the indirect negative effect overwhelms the direct positive effect, consistent with the magnitude of the coefficients.

Net migration of a province and its neighbors overall has a positive and statistically significant impact on that province's business growth. For non-adjacent provinces, a 128% increase in net migration of a province is associated with a one firm increase per square kilometer in the province itself, which is not economically significant. And while entrepreneurship growth of a province is positively affected by its retail sales, it is negatively affected by its neighbors' retail sales. Both effects are economically significant which provides evidence of a highly competitive inter-provincial market in Vietnam. Looking closer into the total effect of nonadjacent provinces, an 18 million VND increase in retail sales per person is related to a one firm increase per square kilometer. In short, population and retail sales have significant effects on firm growth in non-adjacent provinces.

With respect to the H factor, the total effect for the variable College is only statistically significant for the non-adjacent provinces suggesting that overall, an increase in highly educated population will lead to business growth in poorer regions. An increase by 0.335% of the percentage of the population in college is linked to a one firm increase per square kilometer, which is not economically significant. The total effects of the variables High and Primsec are economically significant for only the second dataset where they indicate that an increase in either the number of high school students or the number of primary and secondary school students is linked to an increase in the number of enterprises. Overall, the effect of the human factor is statistically but not economically significant for the non-adjacent provinces. Again, though, one should not over-interpret these human capital results, given the weakness of the measures themselves.

The K factor overall shows the negative effect on the growth of the non-adjacent provinces. The estimator for FDI and Revenue is not economically significant across all regions. The coefficient of production of cereal is mostly negative and statistically significant in both its direct and indirect effects, except the direct effect for Hanoi and Ho Chi Minh regions where it is positive. In non-adjacent provinces, the total effect of cereal production on growth is economically significant where an increase in the production of cereal by 588 kilograms per person is related to a decrease in the number of firms by one firm per square kilometer. This negative relationship suggests that the cereal fields take up a large area in a province which reduces the area for firms to locate. On the other hand, the estimators of the production of aquaculture and livestock are positive and economically significant for the Large-Core-City provinces. This evidence implies that aquaculture and livestock production, rather than cereal production, are inputs for the food manufacturing firms in Hanoi and Ho Chi Minh regions. In general, the total effect of the physical material inputs is negative and economically significant for non-adjacent provinces.

The effect of factor L is only economically significant for the direct effect in Large-Core-City provinces where an increase of about 20 workers per firm is systematically related to an increase of one firm per square kilometer. Its total effect, however, is not statistically significant.

The effect of the factor F is positive for all regions of Vietnam but not economically significant in poorer regions. The total effect of the number of hospitals is only economically significant for the urbanized provinces where it has the biggest effect. An increase in the average number of hospitals per thousand people by 0.0022 is related to an increase in the number of enterprises per square kilometer by one. For non-adjacent provinces, the direct effect of the variable Hospital is statistically but not economically significant, which indicates that development of facilities in Hanoi and Ho Chi Minh City regions is more likely to have a positive effect on the development of enterprises than other regions.

In comparison with the SEM, the SDM may have some problems due to missing variables as SDM assumes all explanatory variables are included while SEM test only assumes whether there are missing variables with significant spatial correlation. Despite its potential weaknesses in terms of missing explanatory variables, the Spatial Durbin Model above may still be the most useful since it directly accounts for both fixed effects and spatial effects between dependent and independent variables. More importantly, the robustness test for SEM and SDM

(Appendix 3) suggests that the difference in the coefficients of these two models is systematic, and the SDM coefficients are consistent while those of the SEM are not. Other criteria for choosing SDM over SEM or Fixed Effect Model are the Akaike's Information Criterion (AIC) and the Bayesian Information Criterion (BIC) scores (Table IX). The AIC and BIC scores of the SDM are lower than the SEM's and Fixed Effect's In both datasets, which suggest that the Spatial Durbin Model is the most relevant to apply in this case.

Table IX. Akaike's Information Criterion and Bayesian Information Criterion

		Fixed Effect	***SEM***	***SDM***
AIC	*Non-adjacent*	*118.4934*	*-29.4096*	*-135.9761*
	Core-Cities	*554.0513*	*523.6517*	*416.1287*
BIC	*Non-adjacent*	*365.3831*	*35.3484*	*-14.5549*
	Core-Cities	*643.1457*	*568.1989*	*505.2231*

In sum, the effects of the market are found to be the most significant both statistically and economically. The variables that have the strongest effects on entrepreneurship growth in Vietnam are population and retail sales, with the dueling Own and Neighbor role of density highlighting the value of incorporating full spatial effects of both dependent and independent variables. An evidence is also found on the effect of human capital and facilities factors on regional entrepreneurship growth of the regions around the two biggest provinces of Vietnam.

4.3. Implications

The entrepreneurship gap may simply be a reflection of more general income and growth gaps between regions of Vietnam, as the market factor is clearly the most important driver of small enterprise concentrations. In that spirit, the issue of returning migration should also be evaluated. Returning skilled migrants are becoming more important to local government policy as they bring the potential to help build networks, create further links between emigration and immigration provinces, and directly contribute to the development of the province. There are multiple ways to attract return migration back into the original province. The government can create favorable conditions and opportunities for students who immigrate to another province to attend better colleges to return to their provinces. For example, policies towards college students can be implemented where university tuition will be paid for by the government on the condition that students must return to their hometown after they graduated. Agencies which provide job placement services, skills training, livelihood programs, and give employers a database of skilled workers specifically for returning immigrants will help increase their intention to return to their home provinces. Central and provincial government agencies can partner up with non-government organizations to encourage migrants to spend or invest in their homelands through partnership program or assistance in establishing small businesses. Investing in the infrastructure of professional sectors in which migrants have experience and skills on is also a good way to attract returning migrants. For instance, the lack of scientific research institutes and facilities in a province will make it difficult for qualified researchers to find a job back home.

But concrete policy guidance requires more detailed data which more accurately measure the concerned effects on growth. Valuable data in Vietnam are gathered only at a highly-aggregated regional level and only for several years. Research would be more precise using more specific yearly data at a provincial or city level. The most important variables to gather at a provincial level are the mean years of schooling and the number of skilled labor for each province, which are more relevant measures of the effect of human capital on growth. Variables on transportation infrastructure, warehouses, power plants, and other complementary facilities for businesses are also necessary to analyze the effect of infrastructure on entrepreneurship growth. Differences in province-specific enterprise policies should also be quantitatively better understood.

5. CONCLUSION

The main finding in this paper is that in 63 provinces in Vietnam, there is truly a gap between entrepreneurial growth in Hanoi, Ho Chi Minh City, and their adjacent neighbors from that of the non-adjacent provinces in 2005 to 2013. The paper also finds that provincial characteristics significantly affect entrepreneurship prospects in an individual province as well as its neighboring provinces, which is consistent with the foundational endogenous growth model. The market factor M, human capital H and the facilities factor F show the strongest effects on regional business development. Focusing on the more peripheral regions arguably in need of greater attention, the M factor shows the most significant effect on growth.

In terms of spatial spillover effects, the market factor M overall has a positive direct and negative indirect effect on regional firms' growth. The human capital H overall has a positive effect on the percentage of high school students as well as primary and secondary school students on growth in large-city provinces. Nonetheless, this factor shows little effect in sparser provinces. But given the measurement issues underscored above, the authors may well be missing significant human capital results. Facilities factor F and entrepreneurship growth in Hanoi and Ho Chi Minh City regions have a substantial positive relationship while this relationship in non-adjacent provinces is not nearly as strong.

Based on the key results above, the Vietnam government should consider incorporating more province-specific policies to foster the development of regional entrepreneurship growth, creating more jobs and income to lagging regions. Moreover, since there are significant spillover effect and limited resources, the government should focus its resources on a province in the center of the poor regions and let the spatial spillover promote growth in neighboring provinces. Development policies should focus on market factors; specifically, increasing market purchasing power by increasing income, which comes back to developing human capital, creating more training facilities, improving educational quality so that graduates have a better chance of getting well-paid jobs, building better school system and reducing tuition for poor students, especially at the high school level. Additionally, the government should also focus on policies for bringing back migrants to help boost the market size and provide skilled labor source. Providing favorable conditions for job finding, investment opportunities, and better infrastructures are measures that should be considered when analyzing policy for return migration. It is also necessary to collect more data at a provincial level to precisely measure and eventually understand the causes of business growth in Vietnam.

REFERENCES

Anselin, L. 2001. "Spatial Econometrics." In *A Companion to Theoretical Econometrics*, Edited by Badi H. Baltagi, 310-330, Oxford: Blackwell.

Anselin, L. 2002. "Under the Hood Issues in the Specification and Interpretation of Spatial Regression Models." *Agricultural Economics* 27(3): 247-267.

Arauzo-Carod, J. M. 2005. "Determinants of Industrial Location: An application for Catalan municipalities." *Papers in Regional Science* 84(1):105-120.

Armington, C., and Z. A. Acs. 2002. The Determinants of Regional Variation in New Firm Formation. *Regional Studies* 36(1): 33-45.

Audretsch, D. B., and M. Fritsch. 1994. "The Geography of Firm Births in Germany." *Regional Studies* 28(4): 359-365.

Bunten, D., Weiler, S., Thompson, E., and S. Zahran. 2015. "Entrepreneurship, Information, and Growth." *Journal of Regional Science* 55(4): 560-584.

Coughlin, C. C., and E. Segev. 2000. "Location Determinants of New Foreign-Owned Manufacturing Plants." *Journal of Regional Science* 40(2): 323-351.

Dejardin, M., and M. Fritsch. 2011. "Linking Net Entry to Regional Economic Growth." *Small Business Economics* 36(4): 443-460.

Dollar, D. 2002. "Reform, Growth, and Poverty in Vietnam." *World Bank Policy Research Working Paper* No. 2837.

Ellison, G., Glaeser, E. L., and W. R. Kerr. 2007. "What Causes Industry Agglomeration? Evidence from Coagglomeration Patterns." *National Bureau of Economic Research* No. w13068.

Freeman, D. B. 1996. "Doi Moi Policy and the Small-Enterprise Boom in Ho Chi Minh City, Vietnam." *Geographical Review*, 178-197.

Garofoli, G. 1994. "New Firm Formation and Regional Development: The Italian Case." *Regional Studies* 28(4): 381-393.

General Statistical Office (GSO). (Various issues). *Statistical Year Book of Vietnam*. Hanoi: Statistical Publishing House.

Government of Vietnam. 1992. Constitution 1992. Foreign Language Publishing House, Hanoi.

Guesnier, B. 1994. "Regional Variations in New Firm Formation in France." *Regional Studies* 28(4): 347-358.

Hakkala, K., and A. Kokko. 2007. "The State and the Private Sector in Vietnam." *Stockholm: The European Institute of Japanese Studies*.

Hansen, H., Rand, J., and F. Tarp. 2009. "Enterprise Growth and Survival in Vietnam: Does Government Support Matter?. *The Journal of Development Studies* 45(7): 1048-1069.

Hue, T. N. 2015. "Income Gap in Regions of Vietnam." *General Statistical Office Journal*. May 2015 (497).

Klier, T. H. 2006. "Where the Headquarters Are: Location Patterns of Large Public Companies, 1990-2000." *Economic Development Quarterly* 20(2): 117-128.

Le, Q. V., and P. V. Raven. 2015. "Woman Entrepreneurship in Rural Vietnam: Success and Motivational Factors." *The Journal of Developing Areas* 49(2): 57-76.

LeSage, J. P. 2008. "An Introduction to Spatial Econometrics." *Revue d'économie Industrielle*, (123): 19-44.

Lu, J., and Z. Tao. 2009. "Trends and Determinants of China's Industrial Agglomeration." *Journal of Urban Economics* 65(2): 167-180.

Malesky, E., and M. Taussig. 2009. "Out of the Gray: The Impact of Provincial Institutions on Business Formalization in Vietnam." *Journal of East Asian Studies* 9(2): 249-290.

Mankiw, N. G., Romer, D., and D. N. Weil. 1992. "A Contribution to the Empirics of Economic Growth." *Quarterly Journal of Economics* 107(2): 407-437.

Martin, R., and P. Sunley. 1998. "Slow Convergence? The New Endogenous Growth Theory and Regional Development." *Economic geography* 74(3): 201-227.

Meyer, K. E., Tran, Y. T. T., and H. V. Nguyen. 2006. "Doing Business in... Vietnam." *Thunderbird International Business Review* 48(2): 263-290.

Mushinski, D., and S. Weiler. 2002. "A Note on the Geographic Interdependencies of Retail Market Areas." *Journal of Regional Science* 42(1): 75-86.

Mushinski, D., Weiler, S., and B. Widner. 2014. "The Impact of Retail Establishments in Hinterlands on the Export Role of Retail Establishments in Rural Places." *Annals of Regional Science* 52: 469-487.

Oostendorp, R. H., Trung, T. Q., and N. T. Tung. 2009. "The Changing Role of Non-Farm Household Enterprises in Vietnam." *World Development* 37(3): 632-644.

Rand, J., Tarp, F., Dzung, N. H., and D. Q. Vinh. 2007. "Documentation of the Small and Medium Scale Enterprise (SME) Survey in Vietnam for the Year 2002." Online Document. Retrieved May, 10.

Romer, P. M. 1994. "The Origins of Endogenous Growth." *The Journal of Economic Perspectives* 8(1): 3-22.

Santarelli, E., and H. T. Tran. 2012. "Growth of Incumbent Firms and Entrepreneurship in Vietnam." *Growth and Change* 43(4): 638-666.

Sutaria, V., and D. A. Hicks. 2004. "New Firm Formation: Dynamics and Determinants." *The Annals of Regional Science* 38(2): 241-262.

Van Arkadie, B., and R. Mallon. 2003. "Vietnam: A Transition Tiger?" Canberra: ANU Press.

Viton, P. A. 2010. "Notes on Spatial Econometric Models." *City and Regional Planning* 870(03): 9-10.

Walder, A. G., and G. H. Nguyen. 2008. "Ownership, Organization, and Income Inequality: Market Transition in Rural Vietnam." *American Sociological Review* 73(2): 251–269.

Wang, S. W. 2006. "Determinants of new firm formation in Taiwan." *Small Business Economics* 27(4-5): 313-321.

Wennekers, S., and R. Thurik. 1999. "Linking Entrepreneurship and Economic Growth." *Small Business Economics* 13(1): 27-56.

Wiklund, J., Patzelt, H., and D. A. Shepherd. 2009. "Building an Integrative Model of Small Business Growth." *Small Business Economics* 32(4): 351-374.

Appendix 1. Correlation Matrix

	Pop	NetMigra	FDI	Retailsale	Freight	Labor	Revenue	College	High	Primsec	Hospital	Cereal	Aqua	Livestock
Pop	*1.000*													
NetMigra	*-0.017*	*1.000*												
FDI	*-0.001*	*0.112*	*1.000*											
Retailsale	*0.112*	*0.195*	*0.054*	*1.000*										
Freight	*0.313*	*-0.063*	*-0.012*	*-0.016*	*1.000*									
Labor	*0.210*	*0.396*	*0.106*	*-0.064*	*0.273*	*1.000*								
Revenue	*0.164*	*0.249*	*0.281*	*0.489*	*0.077*	*0.345*	*1.000*							
College	*-0.061*	*0.093*	*0.205*	*0.063*	*-0.054*	*-0.026*	*0.307*	*1.000*						
High	*0.112*	*-0.142*	*0.082*	*-0.164*	*0.161*	*-0.001*	*-0.056*	*0.097*	*1.000*					
Primsec	*-0.234*	*-0.100*	*-0.022*	*-0.514*	*-0.231*	*-0.138*	*-0.354*	*0.068*	*0.198*	*1.000*				
Hospital	*-0.572*	*-0.043*	*-0.105*	*-0.216*	*-0.179*	*-0.125*	*-0.245*	*-0.052*	*0.116*	*0.233*	*1.000*			
Cereal	*0.163*	*-0.258*	*-0.103*	*0.114*	*-0.004*	*-0.291*	*-0.026*	*-0.171*	*-0.430*	*-0.221*	*-0.259*	*1.000*		
Aqua	*0.143*	*-0.267*	*0.033*	*0.345*	*-0.031*	*-0.321*	*0.212*	*0.012*	*-0.384*	*-0.286*	*-0.437*	*0.530*	*1.000*	
Livestock	*0.098*	*-0.155*	*-0.083*	*-0.174*	*0.055*	*0.111*	*0.018*	*-0.142*	*0.223*	*-0.287*	*0.111*	*0.064*	*-0.253*	*1.000*

Appendix 2. Variance Inflation Factor of the First Difference Regression

	Sparser-Population		Large-Core-City	
	VIF	1/VIF	VIF	1/VIF
Pop	1.08	0.9242	1.72	0.5799
NetMigra	1.08	0.9247	1.14	0.8787
FDI	1.04	0.9587	1.57	0.6383
Retailsale	1.21	0.8284	1.53	0.6551
Freight	1.19	0.8392	1.95	0.5137
Labor	1.31	0.7652	1.73	0.5774
Revenue	1.17	0.8546	2.11	0.4744
College	1.00	0.9951	1.21	0.8246
High	1.58	0.6311	1.22	0.8183
Primsec	1.54	0.6474	1.11	0.9014
Hospital	1.10	0.9052	1.16	0.8602
Cereal	1.05	0.9534	1.28	0.7815
Aqua	1.05	0.9480	1.33	0.7529
Livestock	1.07	0.9344	1.07	0.9341
Mean VIF	1.18	.	1.44	.

Appendix 3. Hausman Specification Test for SDM and SEM

	SDM	SEM	Difference	S.E.
Pop	0.0184	0.0196	-0.0012	.
NetMigra	-0.0390	-0.0461	0.0071	.
FDI	0.0135	0.0306	-0.0171	.
Retailsale	0.2004	0.0791	0.1213	0.0075
Freight	-0.5373	-0.3742	-0.1632	.
Labor	0.0447	0.0547	-0.0100	.
Revenue	-0.0123	-0.0164	0.0041	.
College	-0.7679	0.6754	-1.4433	.
High	55.6261	19.7997	35.8264	8.4699
Primsec	-6.2542	7.3109	-13.5651	1.6910
Hospital	123.9440	103.1010	20.8429	.
Cereal	0.0014	-0.0019	0.0033	0.0002
Aqua	-0.0027	-0.0038	0.0011	0.0004
Livestock	0.0001	-0.0001	0.0002	.

H_0: Difference in coefficients not systematic.
χ^2 (9) = 181.85.
p = 0.0000.

Appendix 4. Hausman Specification Test for SDM Regressions with Area and Population normalization for the dependent variable

	Area Normalized	*Population Normalized*	*Difference*	*S.E.*
		Main		
NetMigra	0.0076	0.0026	0.0050	0.0010
FDI	-0.0004	-0.0082	0.0077	0.0050
Retailsale	0.0838	0.0752	0.0086	0.0016
Freight	-0.5774	-0.6865	0.1091	0.0441
Labor	0.0004	-0.0036	0.0041	0.0010
Revenue	0.0017	-0.0043	0.0060	0.0005
College	-0.3821	-0.6991	0.3170	0.1282
High	-16.5458	-5.8687	-10.6770	2.0240
Primsec	2.7498	0.2620	2.4878	0.4115
Hospital	26.0817	24.2317	1.8499	4.8095
Cereal	-0.0002	-0.0003	0.0001	0.0001
Aqua	-0.0007	-0.0006	-0.0001	0.0002
Livestock	0.0000	0.0000	0.0000	0.0000
		W_X		
NetMigra	0.0083	0.0193	-0.0110	0.0021
FDI	0.0184	0.0304	-0.0121	0.0100
Retailsale	-0.0615	-0.0339	-0.0277	0.0025
Freight	0.7539	0.4982	0.2558	0.0804
Labor	-0.0046	0.0015	-0.0061	0.0018
Revenue	0.0074	0.0050	0.0024	0.0013
College	0.3003	0.0537	0.2467	0.2136
High	20.0353	18.8486	1.1867	2.5503
Primsec	-1.9990	-4.3292	2.3302	0.4757
Hospital	5.4413	-9.2543	14.6956	8.1716
Cereal	-0.0012	-0.0010	-0.0002	0.0001
Aqua	0.0026	0.0010	0.0016	0.0003
Livestock	0.0000	-0.0001	0.0001	0.0000
Spatial				
ρ	0.3690	0.3913	-0.0222	0.0066
σ^2_e	0.0468	0.0360	0.0108	0.0018

Appendix 4. (Continued)

	Area Normalized	Population Normalized	Difference	S.E.
		Direct		
NetMigra	0.0087	0.0049	0.0037	0.0008
FDI	0.0026	-0.0039	0.0066	0.0059
Retailsale	0.0807	0.0750	0.0057	0.0020
Freight	-0.5360	-0.6764	0.1405	0.0487
Labor	0.0001	-0.0034	0.0036	0.0010
Revenue	0.0026	-0.0039	0.0065	0.0006
College	-0.4083	-0.7644	0.3561	0.1189
High	-15.1744	-4.0171	-11.1573	1.8672
Primsec	2.6460	-0.2507	2.8967	0.3951
Hospital	28.3492	24.7065	3.6427	5.0382
Cereal	-0.0004	-0.0004	0.0001	0.0001
Aqua	-0.0005	-0.0006	0.0001	0.0002
Livestock	0.0000	0.0000	0.0000	0.0000
		Indirect		
NetMigra	0.0166	0.0312	-0.0146	.
FDI	0.0300	0.0443	-0.0143	0.0132
Retailsale	-0.0446	-0.0065	-0.0380	0.0045
Freight	0.7524	0.3072	0.4453	0.0844
Labor	-0.0060	0.0007	-0.0067	0.0023
Revenue	0.0114	0.0047	0.0067	0.0019
College	0.1858	-0.3797	0.5655	0.2538
High	20.2507	24.9325	-4.6818	2.6652
Primsec	-1.5331	-6.4870	4.9539	0.5555
Hospital	20.1855	-1.3151	21.5006	10.8733
Cereal	-0.0019	-0.0017	-0.0002	0.0002
Aqua	0.0034	0.0011	0.0023	0.0003
Livestock	0.0000	-0.0001	0.0001	0.0000
		Total		
NetMigra	0.0252	0.0361	-0.0108	.
FDI	0.0327	0.0404	-0.0077	0.0167
Retailsale	0.0361	0.0684	-0.0324	0.0046
Freight	0.2165	-0.3693	0.5858	0.0804
Labor	-0.0059	-0.0027	-0.0032	0.0026
Revenue	0.0140	0.0008	0.0131	0.0021
College	-0.2225	-1.1441	0.9215	0.3073
High	5.0764	20.9154	-15.8391	2.8456
Primsec	1.1128	-6.7377	7.8505	0.6010
Hospital	48.5347	23.3914	25.1433	12.8837
Cereal	-0.0023	-0.0021	-0.0001	0.0002
Aqua	0.0029	0.0005	0.0024	0.0004
Livestock	0.0000	-0.0001	0.0001	0.0000

H_0: Difference in coefficients not systematic.
χ^2 (34) = 663.56.
p = 0.0000.

Appendix 5. Summary Statistics of the Spatial Durbin Model with Population Normalization of the Dependent Variable

	Without HN, HCM (N = 549)				Non-Adjacent Provinces (N = 423)				HN, HCM and Neighbors (N = 144)			
	Mean	Std. Dev.	Min.	Max.	Mean	Std. Dev.	Min.	Max.	Mean	Std. Dev.	Min.	Max.
Growth (P)	1.549	1.078	0.328	9.582	1.464	1.054	0.328	9.582	2.732	2.922	0.647	15.438
Density	387.976	343.258	35.000	1347.000	327.809	319.843	35.000	1260.000	844.632	791.693	171.000	3732.000
NetMigra	-1.560	8.186	-27.300	74.600	-2.687	5.394	-27.300	36.200	3.105	12.976	-11.800	74.600
FDI	0.248	1.029	0.000	17.479	0.184	0.993	0.000	17.479	0.444	1.053	0.000	9.532
Retailsale	11.977	8.818	1.119	49.319	11.708	8.482	1.119	49.319	15.853	14.413	1.750	79.161
Freight	0.381	0.381	0.021	2.683	0.403	0.425	0.021	2.683	0.288	0.139	0.043	0.651
Labor	40.628	19.277	14.110	150.333	36.088	14.706	14.110	108.276	52.746	24.462	19.884	150.333
Revenue	21.541	19.562	2.603	186.265	17.140	11.325	2.603	85.277	34.777	29.268	3.804	186.265
College	0.026	0.077	0.000	0.539	0.019	0.058	0.000	0.498	0.050	0.112	0.001	0.539
High	0.034	0.009	0.010	0.058	0.034	0.009	0.010	0.058	0.034	0.007	0.013	0.053
Primsec	0.164	0.032	0.090	0.480	0.170	0.033	0.112	0.480	0.144	0.022	0.090	0.200
Hospital	0.013	0.005	0.005	0.033	0.013	0.005	0.007	0.033	0.011	0.003	0.005	0.018
Cereal	553.897	431.225	21.000	2578.800	572.770	447.970	34.800	2578.800	446.900	363.298	11.700	1931.100
Aqua	89.989	116.734	0.537	588.779	100.752	125.356	0.537	588.779	48.871	67.314	4.805	288.230
Livestock	3502.296	1685.352	343.481	9922.283	3121.281	1295.854	343.481	7877.005	4347.670	2364.395	26.378	9922.283

Appendix 5. (Continued)

	Total		Direct Effect		Indirect Effect	
	Nonadjacent	Large-Core	Nonadjacent	Large-Core	Nonadjacent	Large-Core
Density	-0.0026	0.0009	**0.0028	**0.0029	***-0.0093	-0.0020
NetMigra	***0.0275	0.0099	**0.0028	-0.0022	**0.0203	0.0116
FDI	0.0374	0.1143	0.0033	-0.0168	0.0353	0.1308
Retailsale	***0.0804	***0.0763	***0.0641	***0.0986	0.0157	-0.0225
Freight	0.1305	***-3.4850	***-0.4420	***-2.2669	*0.5881	-1.0224
Labor	-0.0009	0.0011	-0.0044	-0.0023	0.0026	0.0015
Revenue	-0.0140	-0.0016	**-0.0088	**-0.0059	-0.0052	0.0043
College	2.0451	-1.6637	0.6699	**-1.2234	1.3701	-0.4380
High	***27.3651	8.6824	-4.3357	-32.4746	***30.9979	*40.5659
Primsec	***-6.2293	-6.1608	-0.4741	**24.2414	***-5.7404	*-29.3068
Hospital	41.9361	62.0123	19.6913	**71.0781	18.1705	-0.0355
Cereal	***-0.0021	**-0.0028	-0.0006*	***-0.0044	**-0.0015	0.0016
Aqua	0.0001	*** 0.0554	0.0003	0.0079	-0.0001	***0.0485
Livestock	-0.0001	* 0.0002	-0.0001	***0.0001	-0.0001	0.0004
ρ	***0.4364	*-0.3014	***0.4364	*-0.3014	***0.4364	*-0.3014
σ^2_e	***0.0275	*** 0.0851	***0.0275	***0.0851	***0.0275	***0.0851
N	423	144	423	144	423	144
Provinces N	47	16	47	16	47	16
Mean FE	1.4105	0.9797	1.4105	0.9797	1.4105	0.9797
R^2 Within	0.8758	0.9524	0.8758	0.9524	0.8758	0.9524
R^2 Between	0.3234	0.7646	0.3234	0.7646	0.3234	0.7646
R^2 Overall	0.3628	0.7534	0.3628	0.7534	0.3628	0.7534

Note: The dependent variable is the average number of enterprises per thousand people, 2005-2013. Coefficients with one, two, and three stars are significant at 10 percent, 5 percent, and 1 percent levels, respectively, using a two-tailed test. The "rho" is the coefficient of the spatially lagged dependent variable. The Direct Effect is the effect of the original province's Independent variables on that province's growth taking feedback effect into account. The Indirect Effect is the effect of the independent variables of a province's neighbors on that province's growth taking feedback effect into account. HN, Hanoi; HCM, Ho Chi Minh.