Household business index in Vietnam: a small area estimation method

Cuong Nguyen and Trang Do and Thu Phung and Tung Phung

20. January 2012

Household Business Index in Vietnam:

A Small Area Estimation Method

Nguyen Viet Cuong
Do Thu Trang
Phung Thi Thanh Thu
Phung Duc Tung[1]

Abstract

This study aims to estimate of Business or Formality Index (FI) for business households at provincial and district levels in the rural areas of Vietnam. In order to construct the FI, we implemented a survey of 900 business households in three provinces in Vietnam. Then we estimate the FI for these 900 sampled households. Next, we use a small area estimation method to combine this business household survey and the 2009 Vietnam Population and Housing Census to estimate the FI for all provinces and districts in rural Vietnam. The FI at the national level is estimated to be at 62.7 (according to the score range 1 – 100, the higher the score is, the better the index reflects business environment). The FI varies more greatly at provincial level, from 57.3 to 69.5. The FI for households tends to be higher at more developed provinces and lower at less developed provinces.

Keyword: household business, business environment, Formality Index, household business survey, small area estimation, Vietnam.

JEL Classification: M13; M20; M21.

[1] We would like to acknowledge various researchers including but not limiting to Mrs. Tran Thi Quynh Chi, Mr. Ho Van Bao, Mr. Le Dang Trung, Mrs. Vu Thi Bich Ngoc, Ms. Nguyen Thi Nhung for their contributions in coordinating the research, developing database and providing necessary literature and documents. We are also deeply grateful to the Institute of Policies and Strategies for Agriculture and Rural Development (IPSARD) and the Spanish Development Agency for their funding and support for the achievement of the study. My appreciation also goes to the Provincial Statistics Office of Phu Tho, Dak Lak and Can Tho for assisting us in collecting the data for the Household Business Survey 2011.

Contact author: Nguyen Viet Cuong, email: c_nguyenviet@yahoo.com

1. Introduction

Following the Economic Reform in 1986, Vietnam has achieved remarkable economic growth and poverty reduction in the past two decades. GDP growth rates stands at 7% in the past decade. Poverty rate declines from 58% in 1993 to 37% in 1998, 14% in 2008. Household business contributes significantly to the economic development.[2] Household business not only provides jobs, increases income and GDP but also creates a wide business network to remote areas where enterprises are not developed. A large number of studies show that non-agricultural production household activities contribute significantly to poverty reduction (Lanjouw 1995: Lanjouw 1998: Ruben and Van den Berg 2001). Several studies also indicate that the non-farm sector has helped improve the living standard in Vietnam (Van der Walle 1994, Pham et all 2010).

In the long term, the economy will witness the gradual shift from informal sector to formal sector. However, in the difficult economic setting of the country, household business in the informal sector still plays a key role. Household businesses help reduce the negative impact of economic shocks on household economic condition. Globalization can bring about opportunities together with economic risk (Easterly and Kraay 2000; Winters et al. 2004). When losing formal employment, labors can find job opportunities from household business. Nguyen (2010) demonstrates that some low-income laborers in enterprise sector have shifted to household business area. Farm households when encountering natural or economic shocks also tend to shift to non-farm activities.

The Survey on non-farm individual business establishments by the General Statistic Office (GSO) in 2004 suggests that there were 2.9 million non-farm business establishments across the country creating nearly 5 million jobs. The number of business establishments as well as their employees increases year-after-year. By 2010, there are 4.1 million business establishments with 7.4 million jobs.

However, the growth of such establishments is not as significant as that of enterprises in terms of capital and employment. The former often has small capital and

[2] We use the term Household Business according to Circular 43/2010/ND-CP of the Government. Household business in this case is household or business establishment that carries out non-farm activities. In the report, we use the term 'household business' and 'business establishment' with similar meanings.

employment size which hardly changes after year. On average, each establishment has only 2 workers. The proportion of registered establishments also remains quite low, at only 27.5% in 2007[3]. The business environment for such establishments is generally difficult (Phước Hà, 2006; IRC and IPSARD 2011). Local authorities tend to pay more attention to enterprises since this sector can contribute more to the local revenues through paying taxes (IRC and IPSARD 2011). Besides, employees working for business individual establishments do not receive adequate vocational training and social insurance.

Despite numerous studies on business environment, most of the studies mainly focus on the general environment of a locality or that of enterprises (Freeman et al. 2005; CAP 2007). There is a missing gap on household business, especially rural ones. Among previous studies on enterprises' business environment, a notable one is the research on developing Provincial Competitiveness Index (PCI) conducted by the Vietnam Chamber of Commerce and Industry (VCCI) with the funding of the USAID's Enhancing Vietnam Competiveness Project. Since its first publication in 2005, PCI has become an important tool to measure and evaluate the business environment of the private sector in 63 provinces and cities in Vietnam. It is also helpful for local authorities in identifying obstacles to enterprises' business, especially those related to institution. Nevertheless, PCI is developed only for enterprises or the formal sector. Its scope does not involve household business.

Therefore, in this study, we will develop a Business Index or Formality Index (FI) to examine the business environment of business households in rural Vietnam. Using the FI, local governments can monitor its business environment and have relevant measures to support the business sector for households.

A major challenge is how to estimate the FI for all provinces and districts of the country. In order to estimate the FI for all provinces and cities, we would need a large sample of households (there are 63 provinces and cities in Vietnam). The sample size is much larger if we want to estimates the FI for all the districts (the total number of districts is 690). The large sample of households requires a substantial survey cost.

Thus, in this study, we will use a 'small area estimation' method that is developed by Elberts et al (2002, 2003) to combine a Household Business Survey in 2011 and the Vietnam Population and Housing Census in 2009 to estimate the FI for all

[3] Accroding to statistics from the Survey on Economic, Administrative and Professional Establishments 2002 and 2007

provinces, cities and districts in rural Vietnam. The information from this study can be useful for local policymakers in monitoring and evaluating changes in the households' business environment and have their policies favourable for household business.

In this study, we use two main data sets. The first is from the Household Business Survey in 2011 which was collected by the Indochina Research and Consulting (IRC) company in Vietnam in collaboration with the Institute of Policy and Strategy for Agriculture and Rural Development (IPSARD), Vietnam Ministry of Agriculture and Rural Development. The survey collected data of 900 non-farm business households on their business performance as well as business environment. The second data set is the 15 percent sample of the Vietnam Population and Housing Census (VPHC) in 2009 which was implemented by the General Statistical Office of Vietnam (GSO).

This paper is structured in five sections. The second section introduces the dataset as well as the characteristics and business environment of household business. The third section presents the methodology to construct the Formality Index of households' business environment. The forth section present the estimation of the Index at province and district level. Finally, fifth section concludes and proposes policy recommendations.

2. Data sources and descriptive analysis

As mentioned, in this study, we implemented a Household Business Survey in 2011 in rural areas in three provinces of Phu THo, Dak Lak and Can Tho. This survey contains information on household business performance and the business environment in rural areas. It is aimed at measuring the impacts of such elements on the establishment and development of household business as well as on the households' shift from informal to formally registered establishments. The survey was conducted by IRC, IPSARD and the Provincial Statistics Office of Phu Tho, Dak Lak and Can Tho in September 2011.

The main content of the questionnaires includes information related to the characteristics of household business such as business sectors, location, general information on the business owners, employment, assets, capital, revenue, profit, owners' evaluation of their performance as well as information on managers.

Meanwhile, general information on business environment includes access to land, capital, employment, production inputs, and output market, business registration procedures, formal and informal charges in paying taxes and fees, information on infrastructure and business supporting services, information access and transparency, households' assessment on local support in input access, administrative procedures as well as client dispute settlement.

Random sampling is applied in three stages in each province. The first stage is to select district, the second stage is to select communes whereas the third stage is to select household business. With reference to findings from the pilot surveys in Bac Ninh and Phu Tho in 2010 and the budget for this survey, we select a sample of 900 household business from 90 communes of 16 districts. In the first stage, in each province, 6 districts are selected. During the second stage, 5 communes in each selected province are picked. And in the third stage, among selected communes in the second stage, in each commune we choose 10 formal households along with 5 others as subtitutes in order to ensure the completion of 900 questionnaires.

Table 1: Household business size by provinces/business sectors

	Total	Male workers	Female workers	Externally-hired workers	Unpaid workers
Total	2,1	1,7	1,2	0,5	1,5
By province					
Phu Tho	2,2	2,0	1,1	0,7	1,5
Dak Lak	1,8	1,5	1,2	0,4	1,4
Can Tho	2,0	1,4	1,2	0,3	1,7
By business sector					
Industry and construction	3,0	2,8	1,4	1,5	1,5
Transportation, warehouse	1,6	1,3	1,0	0,2	1,3
Commerce and Service	1,8	1,3	1,1	0,2	1,5

Source: Household Business Survey 2011

Table 1 shows that the total number of workers and the status of employment utilization in household business by three provinces and three main sectors as well as the size of human resource for rural household business remains quite weak since most of the households are small-slaced and employ manual labourers. The Table demonstrates that each household has an average of 2.05 workers among whom, the number of males usually outweighs that of females. This pattern is still witnessed when it comes to service sector where the rate of male over female workers is 1.34 over 1.12. The majority of household business employ their family members or unpaid workers.

Externally-hired workers account for only an insignificant proportion (approximately 25%).

Table 2 shows capital inclusive of household business' owner equity and loans in three provinces. The average size of one household business (exclusive of fixed assets) by September 1st 2011 remains small at around 124.1 million VNDs. Although industry and construction sector employs the highest number of workers and transportation and warehouse sector employs the lowest number, the biggest share of capital is allocated to the latter sector (with the total capital and fixed asset is 389.3 million VNDs), which is almost 3 times higher than that to the former sector (131.1 million VNDs). This suggests that the former sector requires lower capital but creates more jobs than the latter sector

Table 2: Owners' equity and loans classified by sectors/provinces

(million VNDs)	Total	Industry and Construction	Transportation and warehouse	Commerce and Services	Phu Tho	Dak Lak	Can Tho
Total capital and fixed assets	202,1	131,1	389,3	205,4	183,9	288,3	118,6
Total capital (exclusive of fixed assets)	124,1	76,7	210,4	130,4	111,4	178,1	75,8
Loans	25,2	13,4	23,0	29,6	23,9	35,4	13,2
Owners' equity	93,1	55,4	177,8	97,3	80,7	139,1	59,2

Source: Household Business Survey 2011

Table 3 provides information on monthly pre-tax revenue and profit in 2011. Accordingly, Dak Lak had the biggest profit and revenue. Transportation and warehouse is the sector with the greatest profit (5 million VNDs per month) while Commerce and Services sector has biggest revenue (40.3 million VNDs per month. However, the net profit of the latter is only equivalent to that of the Industry and Construction sector and even lower than that of Transportation and Warehouse sector.

Table 3: Revenue and Profit

	After-tax profit 2011 (*million VNDs)*	Revenue (*million VNDs)*	Profit/ Revenue (%)
Total	12,8	32,1	39,9
By province			
Phu Tho	6,8	18,7	36,2
Dak Lak	28,0	68,4	40,9
Can Tho	7,8	15,9	49,3
By sector			
Industry and Construction	6,2	13,9	44,5

	After-tax profit 2011 (*million VNDs)*	Revenue (*million VNDs)*	Profit/ Revenue (%)
Transportation and Warehouse	9,2	16,1	57,2
Commerce and Services	14,7	40,3	36,6

Source: Household Business Survey 2011

The average ratio between profit and revenue is about 40%. The highest ratio of 57.2% is reported for Transportation and Warehouse, which means that this sector generates higher profit even though it requires larger capital than other sectors. In general, this ratio has reflected a positive picture of household businesses in 2011.

3. Construction of the Formality Index of bousehold business

3.1. Methodology

In this chapter, we present a methodology to develop the Formality Index (FI). The FI is a composite index to reflect different elements of the business environment of household business. There are several ways to classify the business environment of household businesses as well as of enterprises. The business environment can be grouped into two general categories: (i) micro environment inclusive of internal characteristics of household businesses such as capital, employment, land, and capacity and other external factors such as inputs and outputs; and (ii) macro environment inclusive of factors which household businesses cannot exert influence on such as infrastructure, policies and laws (Kotler (1985), Porter (1998) and Parker (2009)). The environment can also be classified into the environment inside the business establishment and that outside the establishment. In this study, we consider factors of the external environment which might have direct or indirect effects on the performance of household businesses. Such factors include micro factors like access to land, capital, and inputs as well as macro ones such as infrastructure and support of relevant authorities.

Two issues arising in the development of the FI or any composite index such as HDI or stock index are identifying the components of the index and estimating the

weights of such components. As different factors have dissimilar effects on household businesses' performance, they will be assigned different weights in the index.

The development of the FI for administrative units such as provinces and districts can be summarized as followed:

- Stage 1: Identify and quantify the components of the business environment of household businesses. These components are estimated for household businesses based on the information on business environment from the Household Business Survey.
- Stage 2: Estimate the weights of each component in the FI. Components which have bigger impacts on household businesses' performance will be assigned higher weights.
- Stage 3: Estimating the FI for each locality based on the weights and values of its components.

Hereinafter we will give a detailed presentation on the FI estimating methodology in three stages.

Stage 1: Identify and quantify components reflecting the business environment of household businesses

This stage includes three steps:

Step 1.1 Identify component indexes

The FI for household business is a composite index representing many aspects which can influence the business environment of household business including both micro and macro environment. An difference between micro and macro environment is that the micro environment includes factors which have direct impacts on household business and household businesses can somehow exert their influence on whereas, the macro environment includes factors with more indirect impacts and not influenced by household businesses' activities

After reviewing relevant literature on business environment such as Kotler (1985), Porter (1998) and Parker (2009) as well as considering the our ability to collect data, we have identified 9 component reflecting the business environment of household business as followed (marked I1 to I9):

1. Land access (I1)

2. Capital access (I2)
3. Labour source and production/commercial input access (I3)
4. Output market access (consumers) (I4)
5. Infrastructure access (traffic, electricity, water, Internet and information) (I5)
6. Market entry index (business registration) (I6)
7. Informal charges (I7)
8. Level of legal and regulatory information transparency (I8)
9. Support of authorities to the growth of household business (I9)

Component indexes are marked I^k with k ranging from 1 to 9. It should be noted that these components are factors reflecting the micro and macro business environment and affecting the performance of household business. They do not include households' inherent elements[4].

Step 1.2. Identify the sub-components of the 9 component indexes

How to measure component indexes is an issue. In order to ensure each component to reflect multiple aspects of the business environment, it will be developed from sub-component indexes. In other words, such component is also composite index which can not be directly measured but needs measuring by other sub-components. Specifically, each component can be comprised by 3 to 10 sub-components. The list of sub-components will be provided in Annex A1. Similarly to components, sub-components also reflect aspects of the business environment. They do not include inherent characteristics of households. These sub-components are developed through households' assessments on respective sub-components in their locality. Through Survey on business establishments, we can acknowledge such assessments.

Sub-component indexes are marked S_j. Information on sub-component indexes are collected directly from the questionnaires. Each sub-component is assigned with ascending scores in which lower score indicates the sub-component's negative impact on household business while higher score indicates positive impact. For instance, for a

[4] As we have already presented, components of FI are not similar to those of PCI. PCI components are not fixed over years. In 2010, the components of PCI are (i) market entry ; (ii) land access ; (iii) transparency ; (iv) time cost ; (v) informal charges ; (vi) activeness ; (vii) Enterprises supporting services ; (viii) Training for workers ; (ix) legal institutions.

sub-component assigned score 1 to 3, higher score indicates the positive impact that sub-component has on the performance of household business. Every household business has their score on each sub-component, for example, S_i^j is the score of household i for sub-component S^j.

After assigning sub-component indexes with relevant score range, it is necessary to standardize these indexes to make them comparable. We use the standardization following VCCI score range (VCCI, 2009). Specifically, the score of S_j of household I on 1 to 100 range is calculated as:

$$S_{i(10)}^j = 99\left(\frac{S_i^j - S_{min}^j}{S_{max}^j - S_{min}^j}\right) + 1 \quad . \qquad (1)$$

While $S_{\min}^j$, $S_{\max}^j$ are the minimum and maximum value of S_j in the sample.

Step 1.3: Scores for sub-component indexes

In this step, the component index is calculated by summing all sub-component indexes. For instance, the component index I_k household *i* is comprised by *n* sub-component indexes S_j calculated as followed:

$$\bar{I}_i^k = \frac{1}{n}\sum_{j=1}^{n} S_{i(10)}^j \qquad (2)$$

Stage 2: estimate the weights of component indexes comprising the FI

Weights of component indexes replicate the level of their influence on the business activities of household business. A component with low weight is the one which does not have significant relevance to households' business activities and vice versa. The first step in this stage is to identify the level of factors' influences on households' business activities.

The following components are selected: profit, revenue growth, and households' assessment on their business performance. Such components are marked Y_1, Y_2 and Y_3. Subsequently, we will run regression of these dependent variables on 9 components (explanatory variables). The three following regression models will be conducted:

$$Y_i^1 = \alpha^1 + \sum_{k=1}^{9} \bar{I}_i^k \beta_k^1 + \varepsilon_i^1, \qquad (3)$$

$$Y_i^2 = \alpha^2 + \sum_{k=1}^{9} \bar{I}_i^k \beta_k^2 + \varepsilon_i^2, \quad (4)$$

$$Y_i^3 = \alpha^3 + \sum_{k=1}^{9} \bar{I}_i^k \beta_k^3 + \varepsilon_i^3, \quad (5)$$

Weights for each component will be calculated from regression findings and represent the importance of respective component's effect on profit, growth and households' growth assessment (Y_1, Y_2 and Y_3).

$$\hat{W}_k^{Y_1} = \frac{\hat{\beta}_k^1}{\sum_{k=1}^{9} \hat{\beta}_k^1}, \quad (6)$$

$$\hat{W}_k^{Y_2} = \frac{\hat{\beta}_k^2}{\sum_{k=1}^{9} \hat{\beta}_k^2}, \quad (7)$$

$$\hat{W}_k^{Y_3} = \frac{\hat{\beta}_k^3}{\sum_{k=1}^{9} \hat{\beta}_k^3}, \quad (8)$$

The final weight of each component is the average sum of three afore-mentioned weights.

$$\hat{W}_k = \frac{\hat{W}_k^{Y_1} + \hat{W}_k^{Y_2} + \hat{W}_k^{Y_3}}{3}, \quad (9)$$

Stage 3:Calculate FI for each household business

For each household business, we can calculate the FI reflecting the household's business environment based on the weights and values of 9 component indexes following this formula:

$$\hat{FI}_i = \sum_{k=1}^{9} \bar{I}_i^k \hat{W}_k, \quad (10)$$

The index to evaluate the competitiveness of business environment of household business for each locality (district, province, and city) is the average sum of local $\hat{FI}_i$

$$P\hat{FI} = \frac{1}{m} \sum_{i=1}^{m} \hat{FI}_i, \quad (11)$$

While m is the number of household businesses in the locality. $\hat{PFI}$ varies from 1 to 10 with higher value indicates more favourable conditions of the locality for the growth of household business.

3.2. Analysis

Table 4 presents the average value of 9 component indexes of provinces and districts in Household Business Survey 2011. Component indexes of Dak Lak tend to be lower than those of Phu Tho and Can Tho. In the two latter provinces, access to land, labour sources and other inputs are relatively easy. There is also a possibility that household business does not have demand for large land area nor labour. Therefore, finding a medium-scaled site as well as an insignificant amout of inputs and labour is not a big challenge for such household business.

Table 4: Component indexes of the PFI

Province	Land access (I1)	Capital access (I2)	Labour access (I3)	Output market (I4)	Infrastructure access (I5)	Market entry (I6)	Informal charges (I7)	Information transparency (I8)	Authorities' support (I9)
Phu Tho	77,4	62,5	81,1	47,7	46,7	52,8	72,2	48,9	46,6
	(1,1)	(1,4)	(0,9)	(0,9)	(0,7)	(0,9)	(1,0)	(1,1)	(1,5)
Dak Lak	76,7	48,9	75,0	40,8	49,1	59,7	66,2	56,7	36,1
	(1,1)	(1,6)	(0,9)	(0,8)	(0,7)	(1,0)	(1,1)	(1,1)	(1,5)
Can Tho	78,1	55,2	78,7	46,1	49,6	58,7	64,5	57,7	45,1
	(1,2)	(1,6)	(0,9)	(0,9)	(0,9)	(0,9)	(1,2)	(0,9)	(1,7)
Total	77,3	57,4	79,0	45,5	47,9	55,8	69,1	52,7	43,4
	(0,7)	(1,0)	(0,6)	(0,6)	(0,4)	(0,6)	(0,6)	(0,7)	(1,0)

Standard errors are in parentheses

Source: Household Business Survey 2011

Authorities' support, output market and infrastructure are underestimated by household business. Those indexes of Dak Lak are lower than those of the other provinces. However, access to and clarification of policy and legal information related to business activities of households are also poorly evaluated. This is possibly because authorities tend to pay more attention to enterprises or formal sector than to household business. Among elements of macro business environment, informal charges are insignificant.

To calculate the weights, we run regression of household's business performance variables on component indexes. Selected performance indicators are the ratio between profit and revenue, revenue growth, and household's assessment on its own performance. The estimation of these indicators is presented in Table 5 in the surveyed provinces and districts. Both indicators of profit/revenue and revenue growth of Can Tho are higher than those of the other provinces.

Table 5: Business performance

Province/District	Ratio between profit and revenue (%) (Y1)	Revenue growth 2010-2011 (%) (Y2)	Proportion of households who assessed that their business was better (%) (Y3)
Phu Tho	36.19	10.89	8.28
	(1.56)	(1.21)	(1.59)
Dak Lak	40.89	6.21	16.31
	(2.33)	(1.22)	(2.14)
Can Tho	49.27	11.86	9.43
	(3.03)	(2.20)	(1.72)
Standard errors are in brackets ()			

Source: Household Business Survey 2011

Regression results are shown in Table 6. In general, component indexes are greater than zero indicating a positive correlation between business performance and the favorable business environment. For indexes with negative coefficients in the regression, respective weights of zero are assigned.

Table 6: Regression results to calculate the weights of component indexes

Explanatory variables	Dependent variables		
	Ratio between profit and revenue (Y1)	Revenue growth 2010-2011 (Y2)	Better business performance (Y3)
Land access (I1)	0,1432**	0,1378***	0,0006
	(0,0669)	(0,0465)	(0,0006)
Capital access (I2)	-0,0338	0,0190	0,0001
	(0,0585)	(0,0376)	(0,0005)
Labour access(I3)	0,3141***	0,0483	-0,0011
	(0,0951)	(0,0784)	(0,0007)
Output market (I4)	0,1344	-0,0241	0,0008
	(0,0921)	(0,0743)	(0,0008)
Infrastructure access (I5)	0,0073	0,0733	0,0015*
	(0,1108)	(0,0858)	(0,0009)
Market entry (I6)	0,0529	-0,0862	0,0010
	(0,1052)	(0,0661)	(0,0008)
Informal charges (I7)	0,2571***	0,0238	-0,0011
	(0,0955)	(0,0675)	(0,0006)
Information transparency (I8)	0,0598	0,0093	0,0014**

Explanatory variables	Dependent variables		
	Ratio between profit and revenue (Y1)	Revenue growth 2010-2011 (Y2)	Better business performance (Y3)
	(0,0743)	(0,0604)	(0,0006)
Authorities' support (I9)	-0,1199*	0,0612	0,0002
	(0,0697)	(0,0482)	(0,0004)
Constant	-11,1598	13,7711	-0,0302
	(12,2828)	(9,7549)	(0,0853)
Number of observations	900	900	900
R-squared	0,041	0,019	0,022
Consistent standard error in brackets () *** p<0.01, ** p<0.05, * p<0.1			

Source: Household Business Survey 2011

Coefficients in the regression models are used to calculate the weights of component indexes following the formula (6), (7), (8), and (9). Table 7 presents the result of weight calculation in which the weight of FI's component indexes are shown in the last column. Other columns show the coefficients of component indexes in the regression of business performance variables. These coefficients have been standardized to comprise a sum of one.

Table 7: Weights of component indexes

	Ratio between profit and revenue (Y1)	Revenue growth 2010-2011 (Y2)	Better business performance (Y3)	Weight in FI
Land access (I1)	0,148	0,370	0,100	0,206
Capital access (I2)	0,000	0,051	0,021	0,024
Labour access (I3)	0,324	0,130	0,000	0,151
Output market (I4)	0,139	0,000	0,138	0,092
Infrastructure access (I5)	0,008	0,197	0,272	0,159
Market entry (I6)	0,055	0,000	0,176	0,077
Informal charges (I7)	0,265	0,064	0,000	0,110
Information transparency (I8)	0,062	0,025	0,255	0,114
Authorities' support (I9)	0,000	0,164	0,037	0,067
Total	1	1	1	1

Source: Household Business Survey 2011

Findings suggest that capital access has the lowest weight while land access has the highest in the FI. The weights of other component indexes are relatively consistent. Elements of the micro business environment such as land access, output and input market play the most important role in the FI. Capital access does not hold a significant role in the FI. Meanwhile, weights macro elements such as infrastructure, informal charges and information transparency are quite high.

Table 8: The FI of household business

Province/Districht	The FI	Standard Error
Province/City		
Phu Tho	62,22	0,42
Dak Lak	60,61	0,48
Can Tho City	62,65	0,49
Phu Tho		
Viet Tri city	60,87	0,84
Phu Tho town	61,77	0,94
Doan Hung District	57,95	1,07
Thanh Ba District	66,92	0,91
Phu Ninh District	64,31	1,03
Lam Thao District	61,24	0,74
Dak Lak province		
Buon Ho	60,68	1,19
Buon Dôn	59,58	1,10
Krong Nang District	67,25	1,32
Krong Pac District	57,66	0,95
Krong Ana	60,50	0,79
Cu Kuin	57,87	0,97
Can Tho City		
Vinh Thanh District	61,49	0,66
Co Do District	58,35	1,05
Phong Dien District	63,27	0,73
Thoi Lai District	68,43	0,86
Total	61,85	0,28

Source: Household Business Survey 2011

Table 8 presents the results of the FI estimation. Can Tho and Phu Tho have quite similar FIs while FI of Phu Tho is slightly lower. At province level, FI ranges from 57 to 68. Thoi Lai district of Can Tho City has the highest FI while Krong Pac district of Dak Lak has the lowest one.

4. Estimation of Formality Index at the district and province level

4.1. Small area estimation method

Household Business survey in this study will be conducted in three provinces. Apparently, we can estimate the FI for three provinces as well as surveyed districts (provided that the number of observations is large enough). A challenge is how we can

estimate the FI for other provinces and districts which are not surveyed. To address this, we propose the 'small area estimation' method developed by Elbers et al (2000, 2003). This method is utilized to estimate poverty measurement indicators at small areas by integrating household surveys (with information on income and consumption) with general surveys (allowing estimation in small areas due to large sample). This method enables the estimation of function relationship between expenditure (or income) and households' characteristics using the household surveys and subsequently, applying the estimated function in the general surveys to estimate the level of expenditure (or income) and living standards of localities at small areas such as communes or districts. This method has been in use to create poverty map for more than 40 countries in the world (Bedi et al, 2007; Bigman and Fofack, 2000).

In this study, we utilize the 'small area estimation' method integrating Household Business Survey 2011 and the 15 percent sample of the Vietnam Population and Housing Census (VPHC) in 2009 to estimate the FI for all provinces and districts in the whole country. Although the Household Business Survey was conducted in three provinces, it collected information on business environment and the FI. Whereas, the 2009 VPHC collected data at district level but was exclusive of information on business environment and the FI.

The 2009 VPHC was collected by the GSO in April 2009. The 2009 VPHC has two components. The first component of the 2009 VPHC collected information on demography and general housing of all households and individuals over the country. The second component is the 15 percent sample of the 2009 VPHC. This sample contains detailed information of 3,692,042 households and 14,177,590 individuals. Information was collected on demography, employment, education, disability, individual migration, households' assets and housing. The most notable feature of the sample is that it has information on whether the household was a non-farm business or not. In this research, we use the sample of non-farm household business. The number of rural non-farm household business is 1,061,782.

FI estimation is described as followed:

Step 1: Use Household Business Survey to calculate FI of household business (with the methodology as described in Section 3).

Step 2: Estimate FI model on variables of household businesses' characteristics using data from Household Business Survey. The model can be displayed as followed:

$$FI_{ic} = \beta_0 + X_c\beta + u_{ic} + \eta_c, \quad (12)$$

While FI_{ic} is the FI of household *i* at commune *c*. X_c is the characteristics variable of commune c where household i reside. These commune variables are calculated from the 2009 VPHC. For example, from the Census, we can calculate the proportion of households having computers or the average household size. Data at commune level are connected with 90 communes selected from the sample of Household Business Survey 2011 and then used for regression (4.1). These commune variables can also include several district variables. The list of these variables and summary statistics are provided in Table A.1 in the Annex.

It should be noted that the explanatory variable X may be the variables of households' characteristics. However, there are no variables on the general characteristics of households in the Household Business Survey 2011 and the 2009 VPHC. While Household Business Survey 2011 collected information on household businesses' characteristics such as assets for business activities, The 2009 VPHC collected information on households' characteristics and domestic assets. Besides, as the 2009 VPHC was conducted two years before the business household survey 2011, FI estimation can be significantly inaccurate if we use household-level variables in regression model (12) (see Nguyen, 2009). The utilization of commune-average variables as the explanatory variables in regression (12) is similar to the methodology of updating poverty map by Nguyen (2011).

Standard error has two parts: household business effect u_{ic} and location effect η_c . This step estimates not only the coefficients in model (12) but also the variance of coefficients and model's standard error. In other words, distribution of coefficients, household business' standard error and location error will be estimated in step 1 using the dataset of the Household Business Survey. It should be noted that Step 1 of building the FI is based on the characteristics of the business environment while Step 2 is based on variables of households' characteristics at commune level.

Step 3: Monte Carlo simulation is applied in the 2009 VPHC to estimate FI. Specifically, in each simulation, values of coefficients and standard error of model (12) will be randomly selected from the distribution estimated in step 1. Values of coefficients, household standard error and location randomly selected in simulation s are $\hat{\beta}^s$, $\hat{u}_{ic}^s$, $\hat{\eta}_c^s$ respectively. The FI of a business establishment in the 2009 VPHC is:

$$FI_{ic}^{s} = X_{ic}\hat{\beta}^{s} + \hat{u}_{ic}^{s} + \hat{\eta}_{c}^{S}. \quad (13)$$

The FI at one locality or location (a province or district for example) is the simple average sum of the FI of all household businesses in that locality/location.

$$\hat{FI}^{s} = \sum_{ic} \hat{FI}_{ic}^{s}. \quad (14)$$

The final FI estimation is the average sum of *s* times of simulations.

$$\hat{FI} = \sum_{s=1}^{S} \hat{FI}^{s}. \quad (15)$$

Finally, variance of $\hat{FI}$ is calculated directly from the value sample after simulation. It should be noted that we estimate the FI for rural areas of provinces and districts.

4.2. Regression model

The first step is to model the FI using the Household Business Survey 2011. The dependent variable is the FI, and explanatory variables are commune and district variables from the 2009 VPHC. The number of observations is 900 households in the Household Business Survey 2011. Table 9 presents the Ordinary Least Squares (OLS) regression and Generalized Least Squares (GLS). According to the 'small area estimation' method by Elbers et al (2002, 2003), OLS regression is run initially to estimate the regression residuals and these residuals are used to estimate GLS enabling the standard error of the variance to change with the explanatory variables.

We run different regression models and found out that FI estimations are quite similar. However, such models yield R-squared of around 0.2, which is not high. Low R-squared can be explained in two ways. Firstly, FI of household business is not highly correlated to household business's performance as well as the household's living standard. The FI is not only related to the business performance but also to many other combined factors such as infrastructure and authorities' support. The changes in such combined factors are difficult to be explained. Secondly, commune-level variables with two-year gap are used (from both the 2009 VPHC and Household Business Survey 2011). Therefore, correlation between explanatory variables and FI variable is not high.

Table 9: Regression log of FI

Explanatory variables	OLS			GLS		
	Coefficient	Standard error	P-value	Coefficient	Standard error	P-value
Constant	3,9145	0,1038	0,0000	3,8055	0,1047	0,0000
Average age of household head	0,0076	0,0018	0,0000	0,0118	0,0018	0,0000
Proportion of households having motorbikes	-0,2792	0,0533	0,0000	-0,2192	0,0523	0,0000
Proportion of households having motorbikes	0,3910	0,0804	0,0000	0,2312	0,0860	0,0073
Proportion of self-employed household members	0,2468	0,0588	0,0000	0,2321	0,0627	0,0002
Proportion of Ethnic Minorities in the communes	0,1664	0,0296	0,0000	0,1960	0,0336	0,0000
Proportion of self-employed household head	0,5848	0,0974	0,0000	0,4875	0,1095	0,0000
Average household size	-0,0910	0,0147	0,0000	-0,0956	0,0161	0,0000
Proportion of people working for their household	-0,8793	0,1329	0,0000	-0,7289	0,1567	0,0000
Proportion of people completing high school	0,5987	0,1016	0,0000	0,4857	0,1068	0,0000
Proportions of households having semi-solid housing	-0,0754	0,0205	0,0003	-0,0782	0,0258	0,0024
Proportions of households having latrines	0,0751	0,0220	0,0007	0,0862	0,0246	0,0005
Average years of schooling of household head's spouse	-0,0410	0,0089	0,0000	-0,0292	0,0095	0,0022
R-squared	0,2056					
Rho (correlation of standard error)	0,10					
Number of Observations	900					

Source: Household Business Survey 2011

It should be noted that the regression model of the FI is not aimed at estimating influential factors to FI but at identifying FI's correlating factors. Therefore, causation analysis is not used in regression in Table 9.

Low R-squared can increase the standard error. However, the standard error can be reduced if we have a large sample in the 2009 VPHC. In such case, FI estimation will become more accurate at province level than at district level.

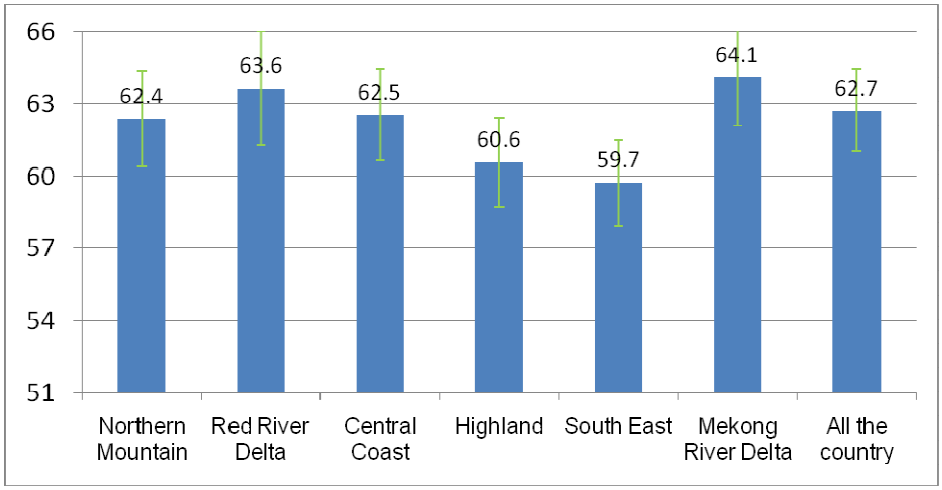
In addition to predicting the FI, we also use the 'small area estimation' method to predict component indexes. These component indexes are also used as dependent variables to run regression on explanatory variables. The regression results are displayed in tables under the Annex.

4.3. Estimation results of the province and district FI

After estimating the FI regression model, we apply this model into the 2009 VPHC to estimate FI at regional, province and district level throughout the whole country. The FI is estimated for rural area and explained for the year 2011. Figure 1 demonstrates the

estimate of the FI for the whole country and six regions. The national FI is 62.5. Among regions, Mekong River Delta has the highest FI followed by Red River Delta. The lowest FI is reported for South East region and Central Highlands. This can be explained by a fact that in South East region, enterprise sector is more developed and household business sector is not as developed as the formal sector.

Figure 1: FI of household business at regional level



Note: Light green lines indicate 90% confidence interval of FI estimation

Source: Household Business Survey 2011

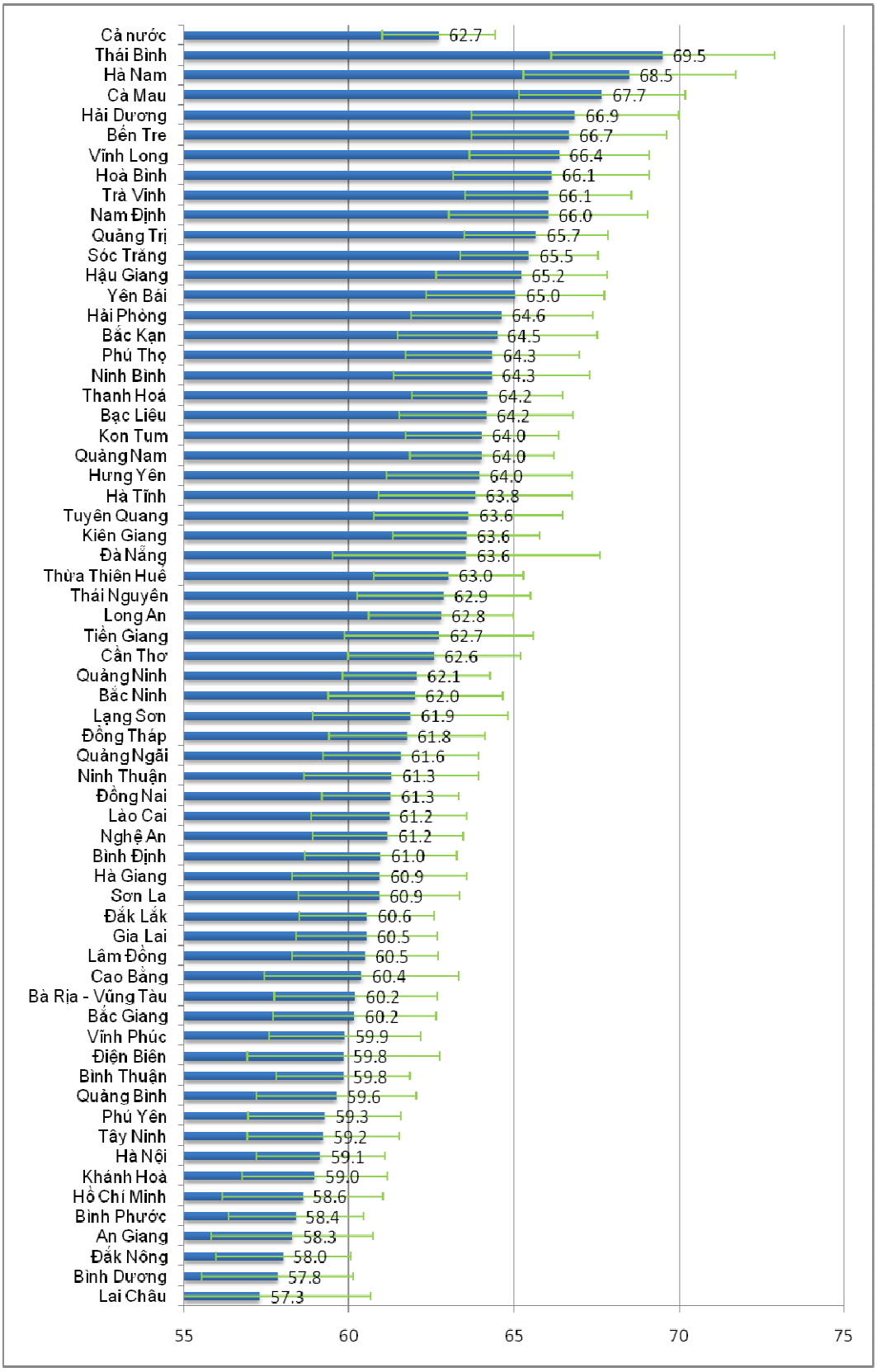
Table 10: FI at regional level

District	FI	Land access (I1)	Capital access (I2)	Labour access (I3)	Output market access (I4)	Infrastructure access (I5)	Market entry (I6)	Informal charges (I7)	Information tranparancy (I8)	Authorities' support (I9)
Northern Uplands	62,4	87,1	49,4	82,4	29,8	48,2	66,6	63,1	40,7	42,1
	(1,2)	(1,9)	(6,3)	(2,1)	(6,7)	(2,9)	(0,7)	(0,9)	(4,3)	(28,0)
Red River Delta	63,6	70,5	64,0	82,1	42,3	45,1	53,0	69,3	53,1	43,3
	(1,4)	(1,9)	(5,0)	(2,1)	(3,8)	(4,7)	(0,9)	(1,4)	(3,3)	(16,2)
Northern Central, Central area	62,5	74,8	48,8	82,4	38,4	52,8	59,6	62,7	56,0	39,7
	(1,1)	(1,1)	(4,5)	(2,1)	(2,9)	(3,3)	(0,7)	(1,0)	(3,0)	(11,6)
Central Highlands	60,6	76,5	45,3	76,9	36,1	50,4	60,8	62,3	48,4	41,3
	(1,1)	(1,4)	(5,6)	(2,0)	(4,1)	(2,7)	(0,7)	(1,0)	(3,8)	(20,0)
Southeast area	59,7	64,9	47,1	78,6	41,9	59,0	56,2	61,3	54,7	39,7
	(1,1)	(1,8)	(3,2)	(2,1)	(3,4)	(3,3)	(0,6)	(1,1)	(3,1)	(11,1)
Mekong Delta	64,H:	74,1	48,7	85,2	47,6	56,5	61,6	66,5	52,8	40,9
	(1,2)	(1,4)	(3,2)	(2,2)	(3,5)	(3,0)	(0,6)	(1,0)	(2,7)	(11,5)
All the country	62,7	75,9	51,8	82,2	39,2	51,0	59,7	64,9	50,9	40,7
	(1,0)	(0,9)	(3,9)	(1,9)	(2,7)	(3,1)	(0,4)	(0,9)	(2,2)	(11,6)

Source: Household Business Survey 2011

Table 10 displays the estimations of component indexes for all regions and the whole country. Some indexes such as output market access vary significantly among regions.

Figure 2: Provincial Formality Index



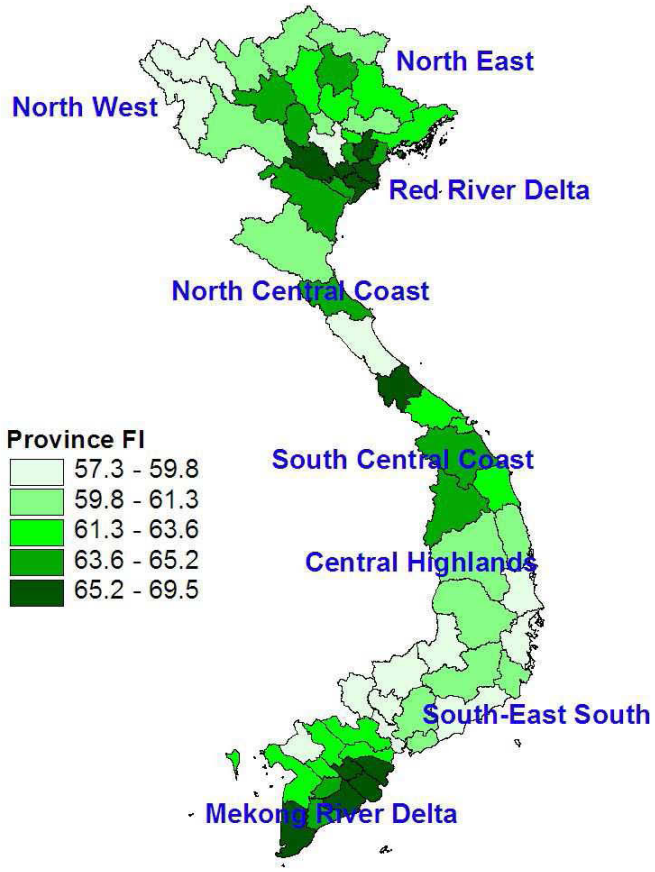
Note: Light green lines indicate 90% confidence interval of FI estimation.

Source: Estimation from Household Business Survey 2011 and VPHC 2009

Figure 2 present the FI estimates at the provincial level and 90% confidence interval of the estimation. FI estimation ranges from 57.3 to 69.5. FIs of Thai Binh, Ha Nam and Ca Mau are 69.5, 68.5 and 58 respectively being the three provinces with highest estimated FI. Meanwhile, Hanoi and HCM City are in the group with lowest FI. In big cities, authorities tend to pay more attention to enterprises. Moreover, land access is more difficult. Another challenge for household business in big cities might come from fierce competition with enterprises or within households.

Figure 3 displays FI among geographical regions. FI varies among provinces in a region.

Figure 3: Map of FI of household business at province level

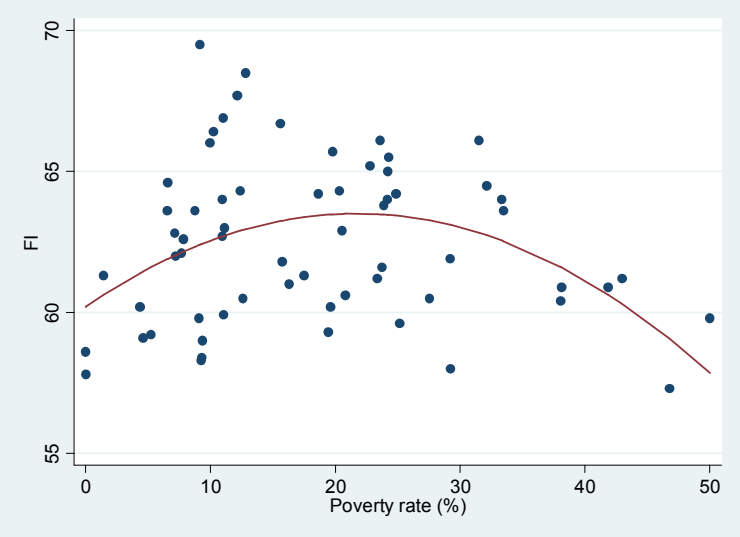


Source: Estimation from Household Business Survey 2011 and VPHC 2009

Figure 4 demonstrates the correlation between FI and the province's poverty incidence. FI tend to be low in province with low poverty rate and tends to increase with growing poverty rates, and to decrease when poverty rates rise significantly. This suggests that business environment of household business is low where enterprises' growth is remarkable. In such regions, labour, capital and production inputs are shifted

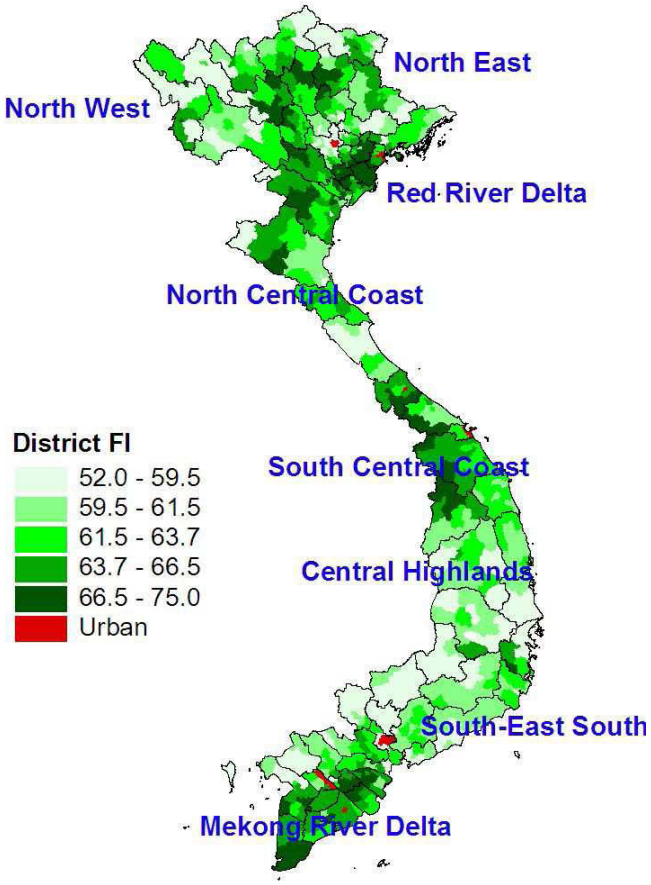
to enterprise sector. On the contrary, in poor regions, business activities encounter many difficulties reflecting an unfavourable business environment for household business.

Figure 4: The FI and poverty rates



Source: Estimation from Household Business Survey 2011 and VPHC 2009

Figure 5: Map of FI of household business



Source: Estimation from Household Business Survey 2011 and VPHC 2009

The 2009 VPHC represents at district level. However, the number of observations of rural non-farm household business is lower, at 1,061,782, which yields large standard error of estimated FIs at district level. Figure 5 shows map of estimated FIs at district level in 2011 used small area estimation method.[5]

5. Conclusion and policy recommendations

Household business has made significant contributions to the country's socio-economic through generating employment, raising income and reducing poverty, especially in rural and remote areas. Household business also contributes to mitigating the negative impacts of economic shocks. Despite its important role, this sector has yet received as much attention from local authorities as enterprise sector.

To develop the FI, this study implemented a survey of 900 business households in three provinces of Phu Tho, Dak Lak and Can Tho in Vietnam. It uses the 'small area estimation' method to estimate the FI for regions, provinces and districts in rural areas throughout the country. The FI is based on the score range from 1 to 100 with higher score reflecting better business environment for household business. The national FI is estimated at 62.7. At the regional level, the FI do not vary remarkable ranging from 59.7 to 64.1. The lowest FIs are reported for South East region and Central Highlands while the highest FIs are estimated for Mekong Delta and Red River Delta.

The FI at provincial levels varies more remarkably, ranging from 57.3 to 69.5. Provinces with the most favourable business environment for household business are Thai Binh, Ca Mau, and Ha Nam whereas those with the least favourable business environment are Lai Chau, Binh Duong and Dak Nong. The FI of household business tend to be higher in medium developed provinces and lower in highly developed and less developed provinces. This can be partly explained by a fact that in highly developed provinces, the formal sector is by far more dominant whereby there are great number of enterprises and fierce competition. The business environment in such provinces might be possibly more favourable towards enterprises than business

[5] In this survey, we estimate FI and component indexes at province and district level. However, in this research, due to limited report length, we do not present the standard errors of component indexes. Interested audience may contact authors to get the results of standard errors of component indexes at province and district level.

households. In poor areas with high poverty rates, business environment of household business is also unfavourable since their market and infrastructure remain less developed.

District FIs reveal that the business environment in Northern districts (except those in North East region), South Central and Mekong Delta are much more favourable than other regions. However, the standard errors of FI estimates are quite large. Therefore, any analysis of the FI should be cautious with standard errors. Overall, FIs at province and district levels can be a helpful tool for local authorities to evaluate and monitor the business environment in their areas.

The Government should have policies to support household business in different business aspects such as capital, access to input, and market. In addition, policies and programs are recommended in rural areas in terms of raising household business owners' awareness of legal issues and of capacity building. Policies targeted business environment need to be aligned with different locality.

References

Bedi, T., Coudouel, A. and Simler, K., (2007), *More Than a Pretty Picture: Using Poverty Maps to Design Better Policies and Interventions*, The World Bank.

Bigman, D. and Fofack, H. (2000), *Geographic Targeting for Poverty Alleviation: Methodology and Applications*. Washington DC: World Bank Regional and Sectoral Studies.

CAP [Trung tâm Tư vấn Chính sách Nông nghiệp] (2007). "Tác động của môi trường đầu tư nông thôn tới sự Figure thành and phát triển của các doanh nghiệp vừa and nhỏ".

Easterly, W. and A. Kraay. 2000. "Small States, Small Problems? Income, Growth, and Volatility in Small States." *World Development* Vol. 28, n11, pp. 2013-27.

Elbers, C., Fujii, T., Lanjouw, P., Ozler, B., Yin, W. (2007), "Poverty Alleviation through Geographic Targeting: How Much Does Disaggregation Help?" *Journal of Development Economics* 83, pp.198-213.

Elbers, C., Lanjouw, J. and Lanjouw, P. (2002), "Micro-Level Estimation of Welfare", Policy Research Working Paper No. WPS 2911. The World Bank.

Elbers, C., Lanjouw, J. and Lanjouw, P. (2003), "Micro-level estimation of poverty and inequality", *Econometrica* 71 (1), pp. 355-364.

Freeman, N., Nguyễn Văn Làn, and Nguyễn Hạnh Nam (2005), "Đăng ký and thành lập doanh nghiệp ở Việt Nam: Con số and thực trang". *Báo cáo Nghiên cứu Chương trình Phát triển Kinh tế Tư nhân MPDF*, Ngân hàng Thế giới.

Kotler, Philip (1985), *Essential Marketing,* Prentice Hall Press.

Lanjouw, J. O. and P. Lanjouw. 1995. "Rural Nonfarm Employment: A Survey." Policy Research Working Paper, 1463. The World Bank: May.

Lanjouw, P. 1998. "Ecuador's Rural Nonfarm Sector as a Route Out of Poverty." Policy Research Working Paper, 1094. The World Bank: March.

Nguyen, V. C. (2009), "Updating Poverty Map without Panel Data: Evidence from Vietnam", *Asian Economic Journal*, Vol. 23(4), 2009, pp. 397-418.

Nguyen, V. C. (2010), "Tác động của Tăng lương tối thiểu lên việc làm ở Việt Nam", Báo cáo cho Bộ Lao động and Thương binh Xã hội.

Nguyen, V. C. (2011), "A Method to Update Poverty Map", *Journal of Development Studies*, accepted.

Parker, Simon (2009), *The Economics* of *Entrepreneurship,* Cambridge University Press.

Pham T., Bui T., Dao T, (2010), "Is Nonfarm Diversification a Way Out of Poverty for Rural Households? Evidence from Vietnam in 1993-2006", PEP-PMMA working papers, Cahiers de recherche PMMA, no. 2010-17.

Phước Hà (2006), "Cở sở sản xuất bị bỏ rơi về chính sách", Vietnam Net, 21/07/2006, http://www.fetp.edu.vn/inthenews/vietnamnet_210706_1.htm;

Porter, Michael E. (1998), *Competitive Strategy: Techniques for Analyzing Industries and Competitors*, Free Press.

Ruben R and M. van den Berg. 2001. "Nonfarm employment and poverty alleviation of rural farm households in Honduras." *World Development* Vol. 29, No. 3, p 549-560, 2001.

Van de Walle, D. 1994. "Rural Poverty in an Emerging Market Economy: Is Diversification into Nonfarm Activities in Rural Viet Nam the Solution?" Policy Research Department. The World Bank, Unpublished Manuscript.

VCCI (2010), "Chỉ số năng lực cạnh tranh cấp tỉnh của Việt Nam năm 2010". *Báo cáo Nghiên cứu chính sách - VNCI,* Số 15, Phòng Thương mại and công nghiệp Việt Nam.

Winters Alan, McCulloch, and Andrew McKay. 2004. "Trade Liberalization and Poverty: The Evidence So Far." *Journal of Economic Literature* Vol. XLII (March 2004).

Appendix

Table A.1: Explanatory variables in the 2009 VPHC

Name of variables	Variable type	Household Business Survey 2011		The 2009 VPHC	
		Mean	Standard variation	Mean	Standard variation
Commune variables					
Proportion of unworking household heads	Continuous	0.141	0.080	0.110	0.082
Proportion of household heads working for the household	Continuous	0.783	0.115	0.824	0.104
Proportion of household heads working for the private sector	Continuous	0.017	0.018	0.013	0.024
Proportion of household heads working for the SOEs	Continuous	0.053	0.061	0.050	0.045
Proportion of household heads	Continuous	0.006	0.033	0.003	0.018
Age of household head	Continuous	46.95	3.27	45.28	4.55
Proportion of male household heads	Continuous	0.791	0.070	0.819	0.081
Schooling years of household heads	Continuous	5.959	1.185	5.565	1.558
Schooling years of household head's spouse	Continuous	4.689	1.161	4.289	1.675
Average household head	Continuous	4.039	0.450	4.103	0.624
Proportion of people without degrees	Continuous	0.381	0.126	0.415	0.157
Proportion of people with primary diploma	Continuous	0.270	0.087	0.266	0.085
Proportion of people with secondary diploma	Continuous	0.219	0.110	0.209	0.113
Proportion of people with higher education diploma (high school, college or university)	Continuous	0.129	0.096	0.111	0.074
Proportion of households with solid wall	Continuous	0.590	0.293	0.558	0.378
Proportion of households with semi-solid wall	Continuous	0.263	0.212	0.245	0.284
Proportion of households with solid roof	Continuous	0.073	0.105	0.094	0.160
Proportion of households with semi-solid roof	Continuous	0.321	0.257	0.410	0.320
Proportion of households with tap water	Continuous	0.088	0.208	0.043	0.138
Proportion of households with clean water (exclusive of those with tap water)	Continuous	0.696	0.307	0.518	0.386
Proportion of households with flush or semi-flush toilet	Continuous	0.334	0.216	0.271	0.250
Proportion of households with toilets	Continuous	0.635	0.219	0.565	0.301
Proportion of household heads with primary diploma	Continuous	0.291	0.116	0.302	0.119
Proportion of household heads with secondary diploma	Continuous	0.259	0.156	0.237	0.157
Proportion of household heads with higher education diploma (high school, college or university)	Continuous	0.158	0.126	0.133	0.092
Proportion of ethnic minority households	Continuous	0.126	0.216	0.365	0.419
Average logarithm of area	Continuous	2.800	0.232	2.723	0.257
Proportions of household with motorbike	Continuous	0.713	0.138	0.645	0.176
Proportions of household with TV	Continuous	0.882	0.074	0.794	0.189
Proportions of household with computer	Continuous	0.053	0.045	0.034	0.042
Proportions of household with refrigerator	Continuous	0.211	0.118	0.148	0.119
Proportions of household with telephone	Continuous	0.403	0.190	0.357	0.198
Proportion of unworking people	Continuous	0.159	0.058	0.126	0.065
Proportion of people working of their household	Continuous	0.510	0.077	0.550	0.079
Proportion of people working for the private sector	Continuous	0.019	0.022	0.012	0.021
Proportion of people working for the state sector	Continuous	0.032	0.030	0.028	0.025
Proportion of people working for the foreign-invested sector	Continuous	0.010	0.031	0.004	0.019
Proportion of children under 15 years old	Continuous	0.270	0.052	0.279	0.064
Proportion of people over 65 years old	Continuous	0.082	0.033	0.080	0.031
Number of enterprises	Continuous	5.756	7.637	5.496	13.357
Total population	Continuous	9456	5034	7321	4766
Total number of households	Continuous	2296	1138	1838	1262

Name of variables	Variable type	Household Business Survey 2011		The 2009 VPHC	
		Mean	Standard variation	Mean	Standard variation
Province variables	Continuous				
Proportion of urban households	Continuous	0.163	0.126	0.136	0.127
Proportion of ethnic minorities households	Continuous	0.105	0.144	0.328	0.359
Total number of households	Continuous	28857	8969	28965	17592
Number of observations	Continuous	900		1061782	

Source: Sample 15% of the Population and Housing