THÈSE DE DOCTORAT

de l'Université de recherche Paris Sciences et Lettres
PSL Research University

Préparée à l'Université Paris-Dauphine

Essays on Vulnerability and Inclusive Development in Developing Asia: A Focus on Vietnam

École Doctorale de Dauphine — ED 543

Spécialité **Sciences économiques**

COMPOSITION DU JURY :

M. Jean-Pierre CLING
INSEE, DIAL
Président du jury

M. Marc LAUTIER
Université Paris 13
Rapporteur

M. Cuong LE-VAN
CNRS
Rapporteur

Mme. Mireille RAZAFINDRAKOTO
IRD-DIAL
Directrice de thèse

M. François ROUBAUD
IRD-DIAL
Co-Directeur de thèse

Soutenue le 27.03.2017
par Thi Thuy Linh LE

Dirigée par **Mme RAZAFINDRAKOTO**

& M. ROUBAUD

DAUPHINE UNIVERSITÉ PARIS | PSL RESEARCH UNIVERSITY PARIS

UNIVERSITÉ PARIS DAUPHINE
ÉCOLE DOCTORALE DE DAUPHINE

Essays on Vulnerability and Inclusive Development in Developing Asia: A Focus on Vietnam

THESE

Pour l'obtention du titre de Docteur en Sciences Économiques

Thi Thuy Linh LE

Jury

Jean-Pierre CLING, INSEE, DIAL	Suffragant
Marc LAUTIER, Université Paris 13	Rapporteur
Cuong LE-VAN, Centre National de la Recherche Scientifique	Rapporteur
Mireille RAZAFINDRAKOTO, Institut de Recherche pour le Développement	Directrice de thèse
François ROUBAUD, Institut de Recherche pour le Développement	Directeur de thèse

Luận án này cháu xin kính tặng ông nội. Bốn chữ "Thư bút minh gia" trên tấm hoành phi ở nhà ông bà, và những câu chuyện về nước Pháp của ông, là duyên phận đã mang cháu đến ngày hôm nay.

Contents

General introduction 9

Chapter 1: Heterogeneity of the informal sector in Vietnam: A *'quali-quanti'* approach 31

Chapter 2: Returns to higher education in Vietnam: Who benefit the most? 99

Chapter 3: Wage differentials between temporary and permanent workers in Asian developing countries 153

Detailed table of contents 216

List of tables 219

List of figures 221

List of acronyms 222

Acknowledgements

I have not managed to find the best words to thank my supervisors, Mireille Razafindrakoto and François Roubaud, for their ceaseless support throughout my PhD research and earlier. With passion, enthusiasm and kindheartedness, they have led or involved in development projects in developing countries including mine. For that, destiny has guided me to their office since my first year of Master study in 2012. Since then, I have been overwhelmed with gratitude by their devoted guidance in all aspects of doing research that one could think of, and their immense tolerance for mistakes that I have made.

My highest appreciation also goes to Professor Le Van Cuong, Marc Lautier for being my 'rapporteurs de these,' and Jean-Pierre Cling for taking part of the Jury. Without their precious comments, this thesis would not have been accomplished in its best state.

I would like to acknowledge Nguyen Huu Thanh Tam's (University of Rouen), Nguyen Huu Chi's (University of Paris 13), and Tran Ngo Thi Minh Tam's (VASS-CAF) direct contributions in parts of my PhD research. Thanh Tam has greatly encouraged me in the last year of my PhD with his research passion, diligence and selfless helps.

I am also indebted to Mariya Aleksynska (ILO) for her valuable expertise and continuous feedbacks within the project on non-standard employment in Asia. Ha Huy Thai (Evry University), Pham Ngoc Sang (Montpellier Business School) and Nguyen Thi Do Hanh (University of Paris-Sorbonne) have my sincere gratitude for their patient mathematical support. I am also grateful to Laure Pasquier-Doumer and Xavier Oudin (IRD-DIAL) for their revision of my work on the informal sector co-authored with Tran Ngo Thi Minh Tam, part of which serves as ingredients for the General Introduction.

I apologize for mentioning just some of the names who have inspired me with brief yet valuable exchanges: Hoang Thuy Quynh, Christophe Nordman; Pham Minh Thai, Vu Hoang Dat, Nguyen Duy Loi (VASS); Pham Hoang Van (Baylor University), Phung Duc Tung (MDRI), Ping Yan (Peiking University), Shirley Johnson-Lans (Vassar College), Nguyen Duc Thanh (VEPR), Myriam de Loenzien (IRD), among others. Axel Demenet (IRD-DIAL) has given me a great help with his experience in defending a PhD thesis in development economics, in the same school and with the same supervisors as mine.

The hospitality of all DIAL members is beyond my expectation. My sincere thanks go to Flore, Philippe, Sandrine, the nicest 'chefs' I have ever met; to Danielle, Anne and Loïc for their unconditional support; to Lisa, Jean-Michelle, Jean-Noël, Delia, Camille, Sarah, Linda, Marin, Björn, Jaime, Marlon, Estelle, Esther, Raphael, Virginie and other *Humans of Enghien* who have made me feel *'comme chez moi'* with their friendliness.

Let me take this chance to thank Paris-Dauphine University for their administrative support and teaching opportunity. I am also grateful to all members of the Vietnamese-French team who participated in the project on the Household Business and Informal Sector in Vietnam

and its sponsors. The ILO in Geneva has kindly granted me with cross-country data and a great chance to participate in their project on Non-standard employment in Asia, this is how chapter 3 was started. I also owe ISVE (International Society of Vietnam Economists) for organizing annual conferences (VEAMs), where I received inspiring ideas and knowledge.

This PhD thesis would not have been realized without the doctoral scholarship of French Government, allocated via the French Embassy in Vietnam and Campus France. Broadly speaking, I am grateful to the generosity of *la France* and French people, with numerous kinds of financial as well as non-financial supports to foreign students.

Finally, I wholeheartedly dedicate this thesis to my beloved family. My grandfather inspired me to pursue doctoral study; but he passed away before he could see I accomplish it. I sincerely thank my parents-in-law for their patience; waiting three years to have a grandson because your daughter-in-law is doing a PhD is a great tolerance. I do not know how to fully express my gratitude to my parents, who have supported me unconditionally all along my way. My thesis is actually theirs. My husband is a great accompanier, without his delicate care, support and encouragement, I may not have finished my PhD in pregnancy.

General Introduction

Six years ago, carrying a bag full of lollipops, my schoolmate and I knocked the door of almost every silk-weaving unit in Van Phuc, a thousand-year-old traditional craft village in Ha Dong, Hanoi. Receiving the lollipops, many children seemed excited and their parents seemed more available to answer the survey questions (at least this was what we felt). But the lollipops did not help in cases of big silk shops in the main street of the village; buying their products plays the role instead. The sharp contradiction between professional and luxurious craft enterprises located in the touristic avenues and depressed household businesses (HBs) in remote hamlets has laid an unforgettable impression on me. While some craftsmen expressed their pride and passion in making original silk products, some of which were even selected for the 1000th anniversary of Thang Long – Hanoi; weaving silk was just to sustain a living for many others. In the aftermath of the economic crisis, some businesses still managed to export their products, whilst most of the interviewees were discouraged and in the worst cases, had to shut down their weaving machines.

The multi-faceted picture of economic growth has become vivid in my heart since this experience in Van Phuc village. Behind luminous achievements in terms of poverty alleviation, education universalization, women empowerment and health care, among others, vulnerability is ubiquitous in Vietnam today. A non-negligible proportion of Vietnamese population are 'excluded' from the triumph melody of economic development in various ways: insufficient education, precarious employment, meager income and so forth. Some of them yet received no direct assistance from the State, and were even deemed to be inexistent. Here comes the motivation for this whole PhD thesis. From my point of view, they are not (only) the 'miserable' who long for society's kindheartedness to sustain a minimum living, but (also) an important engine for country's development thanks to their great potentials which are actually eclipsed by their constraints. Understanding their constraints as well as potentials would thus enable us to position them correctly and help them effectively.

The remainder of General Introduction aims to highlight the importance and relevance of inclusive growth and vulnerability and to present the context in which each research question has been raised. It offers a sketch of Asia and Vietnam in the process of development from a historical perspective. It also provides background information on the studied topics and summarizes the chapters. To some extent, this thesis is dedicated to the underprivileged, and I could say that this is its 'privilege.' As Adam Smith (1723-1790) has said: *"The disposition to admire, and almost to worship, the rich and the powerful, and to despise, or, at least, to neglect persons of poor and mean condition is the great and most universal cause of the corruption of our moral sentiments."*

1. General context

1.1. Asia: Yesterday's glory, today's reemergence and the question of inclusive growth

> *"Merchants and traders travel freely without fear of bandits*
>
> *The jail houses remained empty and people felt no need to lock the gates*
>
> *Bountiful harvests were frequent and ten liters of rice cost only three or four coins."*[1]

This description of a historian about the prosperous Tang dynasty in China reminds us that Asia was once the center of the world. 58 percent of the global economy belonged to Asia prior to the Industrial Revolution in the middle of the 18th century. Asia then gradually lost its dominance to the hand of the West, producing only 15 percent of the world GDP around 1952 (Kohli et al., 2011).

The recent decades have been witnessing the reemergence of Asia, as it accounted for 40 percent of global output and nearly two-third of global growth.[2] The Asian Development Bank has put forward the scenario of an 'Asian century,' with Asia constituting more than half of global GDP and making some three billion additional Asians by 2050, thus regaining the dominant global economic position (Kohli et al., 2011). It is estimated that between 2005 and 2010, 430 million people have been lifted out of poverty, accounting for 93 percent of global poverty reduction during the same period (Chandy & Gertz, 2011). The income in Developing Asia grew 9.4 percent annually over the 2001-2010 decade, reaching roughly 5,000 dollars in purchasing power parity terms in 2010 (Kohli et al., 2011).

Asia is undoubtedly a rising star in the horizon of the globalization era. Yet, according to Noriyuki Suzuki, General Secretary of International Trade Union Confederation – Asia Pacific (ITUC-AP), a core problem sustains:

> *"The biggest challenge facing the Asian and Pacific region and the world is how to institute a truly inclusive and sustainable development paradigm. The present architecture of globalization is largely based on neo-liberal economics, which is at the roots of deepening inequality and persistent poverty across the region and the globe, i.e., the exclusion of the vast humanity of workers from the fruits of the so-called economic surge of Asia and the Pacific."* (Ofreneo, 2013)

Indeed, behind the luminous growth picture, *"520 million people are starving today in Asia, not to mention the millions of people living on the brink of existence."*[3] One-third of the urban population of the Asia-Pacific is living in slums.[4] Ofreneo (2013) also indicates that Asia's growth *"has been both unequal and exclusivist,"* and that the Asian job market is generally precarious. The excluded

[1] [NTDonChina] (2012).

[2] Source: https://www.imf.org/en/News/Articles/2015/09/28/04/53/socar050615a

[3] Decent Work and the Labor Law - 2009-2011, ITUC-AP/LO-FTF Council, 2009.

[4] United Nations Human Settlements Program (UN-Habitat/UNESCAP), 2010.

population from decent work,[5] according to the author, includes informal workers, precarious employment (casual or non-regular) and migrant workers. Generally, these groups of workers are likely subject to instability and inadequacy of protection against abuses and poor working conditions.

Ofreneo (2013) states that the Asian labor markets are 'informalizing.' Particularly, there is a rising trend concerning 'non-standard' or 'irregular workers' in Japan and South Korea; 'floating population' in China and Vietnam; migrants in Singapore, Taiwan, Malaysia; and 'flexies' and 'informals' in Southeast Asia, South Asia and other countries in the region. Informal employment accounts for around 65 percent of non-agricultural employment in Asia-Pacific, according to ILO (2005) estimate. The share of the informal sector, also called vulnerable employment,[6] reaches 50 percent in East Asia, 78 percent in South Asia and 62 percent in Southeast Asia in 2010 (ILO, 2012).

The prevalence of vulnerable employment is widely believed to result from economic policies that encourage deregulation and flexibilization while being apathetic towards organized labor, labor regulations and social protection (Debrah & Smith, 2002). According to Ofreneo (2013), the dominant growth model tends to favor the few with a 'global reach,' whilst ignoring the role in job and wealth creation of other sectors such as domestic industry and agriculture, local artisans and producers, farmers and fisherfolk and the so-called producers of 'non-tradables.' Felipe and Hasan (2006) argue that *"governments across developing Asia must bring the objectives of full, productive, and decent employment to the top of their agendas,"* and emphasize that: *"Unless these three objectives become a fundamental part of the macroeconomic agenda across developing Asia, it is easy to conceive a region, say 25 years from now, which, despite continuous growth, will still harbor most of the world's poor."*

In that context, Ofreneo (2013) calls for a 'social dialogue on economic and labor adjustments', in which he highlights that: *"the biggest dilemma confronting policy makers and legislators in the twin areas of labor law and labor market reforms is how to balance the requirements of business for greater flexibility in the operation, including flexibility in the hiring and firing of workers, with the increasing needs of workers for more, not less, protection given the stability of jobs and incomes in a globalizing economic environment."*

Another important factor deepening the income and job inequality in the region is their development process which favors the highly skilled/educated workers. The World Bank (Gill & Kharas, 2007) attributes the growing inequality in Asia to the growth process driven by economies of scale. Accordingly, within the urban areas, the inequality is caused by higher wage premiums given for skills and talents. Highly qualified workers are more likely to occupy decent work, and simultaneously, tend to achieve higher earnings. Workers without required skills, in contrast, might be trapped in indecent and precarious work, which could be described as *"multitude of insecure people, living bits-and-pieces lives, in and out of short-term jobs, without a narrative of occupational development"* (Guy

[5] According to ILO, "Decent work sums up the aspirations of people in their working lives. It involves opportunities for work that is productive and delivers a fair income, security in the workplace and social protection for families, better prospects for personal development and social integration, freedom for people to express their concerns, organize and participate in the decisions that affect their lives and equality of opportunity and treatment for all women and men." Source: http://www.ilo.org/global/topics/decent-work/lang--en/index.htm

[6] ILO Employment Indicators (2009) defines the 'vulnerables' as the own-account workers and unpaid family workers. The total of these two categories is considered as the size of the informal sector.

Standing, 2011). The ILO (2005) thus underlines the importance of *"making decent work an Asian goal"* in promoting an inclusive and equitable globalization and achieving the Millennium Development Goals (MDGs).

1.2. Vietnam: An ugly duck turns into a swan?

Nowadays, ridiculous stories of the 'subsidy period' (thời bao cấp) such as four people shared a pair of bicycle tyres are still recalled in conversations of Vietnamese. During this period, someone who looked sad were compared with 'losing rice booklet' (mất sổ gạo). Long queues were often seen in front of State department stores, and anything could be used to mark their turn including bricks. People stood with only one foot in prolonged inter-province trains, which may last indefinable due to illegal commercial checking points. Once the ding-dong started, cooperative farmers who had already lifted up their shovel immediately stopped without hoeing up ground. These kinds of story were countless. The centrally planned economy and its underlying ideology have also promoted vicious thinking and habits such as laziness, deceitfulness, bribery and sycophancy. Consequently, in such an agricultural country, many people had to eat rice mixed with tapioca, green banana or sweet potato from time to time. Between 1976 and 1980, Vietnam had to import up to 5.6 million tons of food (Broothroyd & Pham, 2000).

The transition of Vietnam into a market economy has made the above stories 'unimaginable' to most of the Vietnamese youth today. Known as *'Doimoi'* (i.e., Reform, or Renovation), a historic new policy direction had been launched in Vietnam since mid-1980s, including multiple measures such as agricultural de-collectivization, land reforms, reducing subsidies to and increasing autonomy of state-owned enterprises, encouragement of private sector development and trade liberalization (World Bank, 1998). The Reform has transformed Vietnam from one of the world's poorest to a lower middle income country. According to an overview by The World Bank, Vietnam has recorded great improvements in numerous economic as well as social indicators.[7] Per capita income has raised from around 100 U.S. dollars in the 1980s to about 2,100 U.S. dollars in 2015. The proportion of people living in extreme poverty has dropped from about 50 percent in the early 1990s to three percent in 2012. In general, the Vietnamese population today is better educated and has a higher life expectancy than most countries with comparable per capita income.

Nevertheless, there is a downside to *Doimoi* (Boothroyd & Pham, 2000). For households that lack good health, local job opportunities, capital, access to productive land, or adequate accommodation, poverty is not eradicated (ibid). The authors argue that even in the better-off households, some individuals may suffer from adverse effects in terms of narrowed access to and higher costs of formal education, health care, cultural events and a poorer working quality. Actually, a considerable share of the Vietnamese population is still subject to vulnerability. Indecent work prevail: 81 percent of total employment in Vietnam in 2009 were informal (Cling et al., 2014) and 67 percent of wage workers were temporary.[8] Many people could not access education, at any level, including tertiary education (Vu et al., 2012). These people are highly

[7] Source: http://www.worldbank.org/en/country/vietnam/overview [Access date: 8th December, 2016]
[8] Author's calculation based on Labor Force Survey 2007.

vulnerable to economic and health shocks as well as climate change and natural disasters in the absence of governmental assistance.

Thus, understanding the difficulties and discrimination against these disadvantaged groups as well as their potential is meaningful toward a more equitable and inclusive development paradigm of Vietnam. Borrowing Theodore W. Schultz's words, it can be said that: Most of the people in the world are vulnerable, so if we knew the economics of being vulnerable, we would know much of the economics that really matters.[9]

Photos: The 'subsidy period' in Vietnam



"Food stamp: 50 grams"



In a State department store



Transportation



At 'school'

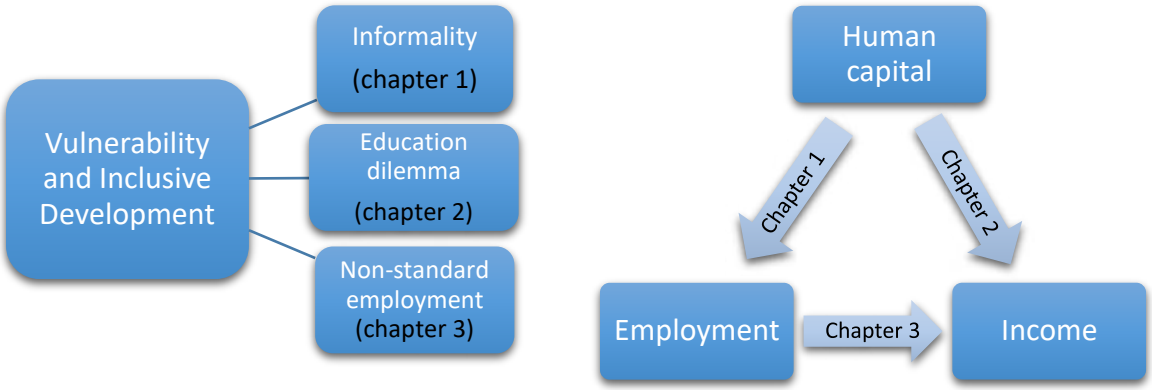
Source: Author's compilation from internet.

[9] His original words: *"Most of the people in the world are poor, so if we knew the economics of being poor, we would know much of the economics that really matters. Most of the world's poor people earn their living from agriculture, so if we knew the economics of agriculture, we would know much of the economics of being poor."* Source: http://www.nobelprize.org/nobel_prizes/economic-sciences/laureates/1979/schultz-lecture.html

2. Chapters' summary and contextualization

This PhD thesis is placed amidst the broad picture of 'exclusivist' growth and vulnerability characterizing Vietnam as well as developing Asia today. The chapters address three important aspects of vulnerability and inclusive development, namely: Informality (chapter 1), Education dilemma (chapter 2) and Non-standard employment (chapter 3). These chapters are complementary as they investigate the interlinkages between *'human capital,' 'employment,'* and *'income.'* Chapter 1 illustrates how human capital, particularly entrepreneurial skill, of Vietnamese HB heads could relate to the functioning and economic performance of their production unit, thus positioning their business differently. Chapter 2 shows the extent to which a greater amount of human capital translates into higher income, by assessing the impact of higher education on individual's earnings and its variation across social backgrounds. Chapter 3 is dedicated to the wage differentials caused by job status in selected Asian countries. Overall, the whole thesis implies that *'human capital,' 'employment,'* and *'income'* are interrelated facets of individual well-being, and that some development phenomena should be analyzed in their heterogeneity.

Graph 1. Thesis in a nutshell: Addressed issues and their interactions



The following paragraphs will present the contexts which lead me to, and exhaustive summaries of the three chapters. While the context section provides necessary definitions, statistics and a historical perspective of the studied subject, the summary paragraph briefly introduces the starting point (i.e., motivation) and the whole procedure through which I have reached the final results, in terms of data, methodology, main findings and policy recommendations.

2.1. Chapter 1: Heterogeneity of the informal sector in Vietnam: A *'quali-quanti'* approach

Context: The HB and informal sector is omnipresent in Vietnam as in most developing countries. Almost every Vietnamese is connected to this sector in everyday life: when you have a beef noodle soup ('pho') on the pavement of a hamlet, buy fruits from an ambulant street vendor,

take a motorbike taxi ride ('xe ôm') or have your hair cut by your retired neighbor. In 2014, the HBs constituted nearly 60 percent of the total number of non-agricultural main jobs in Vietnam.[10] Paradoxically, little effort had been exerted to fully understand and assist this sector from both the researchers and State's sides. Glewwe et al. (2004) has stated that "*it is unlikely that nonfarm household enterprises will play a decisive role […] because their share of the total labor force slowly declined in the 1990s. Their main role is to serve as a temporary source of employment for workers who will eventually find wage work.*"

To improve the understanding of the informal sector's role, characteristics and functioning, the joint project of the General Statistics Office of Vietnam (GSO) and the French development research institute IRD-DIAL (Développement, Institutions & Mondialisation) has been embarked in 2006 (Cling et al., 2010). In 2007, a suitable framework was implemented to measure the informal economy in Vietnam, in accordance with international recommendations and adapted to Vietnamese context. Since then, three waves of Household Business and Informal Sector (HB&IS) survey have been carried out in 2007, 2009 and 2014. Accordingly, the informal sector is defined as *"all private unincorporated enterprises that produce at least some of their goods and services for sale or barter, are not registered (no business license) and are engaged in non-agricultural activities."* In Vietnam, 'unincorporated enterprises' are called 'household businesses,' which either have to be registered (i.e., formal) or not (i.e., informal) depending on their income level (ibid). It is regulated that *"a HB as owned by one Vietnamese citizen, by one group of persons or one individual household may be registered for business at one location only, may employ up to ten (10) employees, shall not have a seal, and shall be liable for its business activities to the full extent of its assets."*[11] If an individual business establishment has more than ten employees and/or at least two business premises, it must be transformed into an enterprise under the Law on Enterprise (Cling et al., 2010). This marks a clear distinction between HBs, even formal ones, and professional enterprises operating in the formal economy. Surveys on employment (Labor Force Survey or LFS) conducted by GSO since 2007 show the preponderance of the informal sector and a slow pace of the formalization process (**Figure 1**).

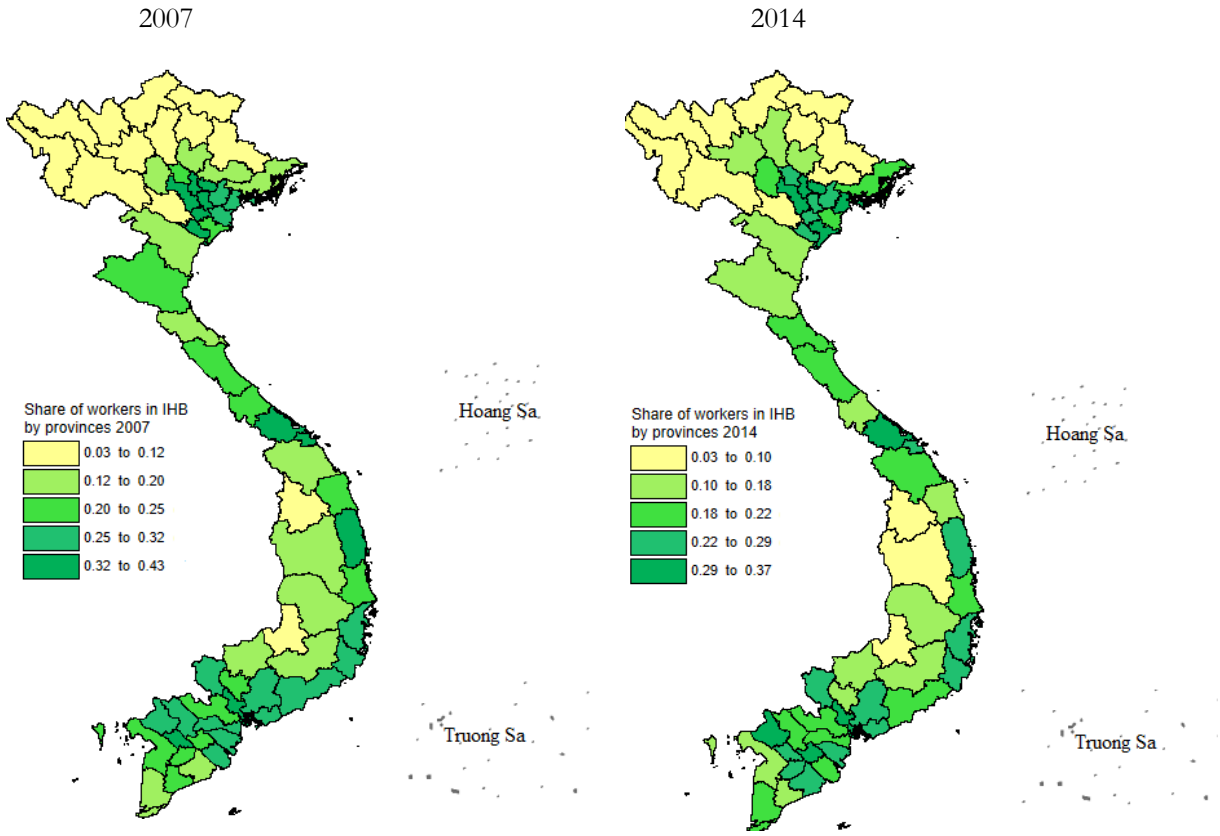
The term *'informal sector'* was first used by Hart (1973) to describe a section of urban labor force constituting unskilled and low-income workers. The distinction between formal and informal activities originally stems from the distinction between wage employment and self-employment. The most recent HB&IS survey (2014) also confirms that in general, this sector is dominated by small-sized informal HBs, mostly own-account workers, operating in precarious conditions and have the lowest income among non-farm sectors. About 90 percent of informal HBs have one or two workers, including the owner, usually with her spouse, another family member or, less often, a hired worker. More than a third of informal HBs operate in unfixed premises, such as the streets, outdoor improvised post, outdoor permanent post, vehicles, customers' home or improvised post in a market. On average, workers in the informal sector attain low levels of formal education, and skills used in informal HBs are mainly gained outside formal education.[12]

[10] Source: HB&IS survey 2014, VASS-CAF & IRD-DIAL team.

[11] Decree N°88 dated 29/08/2006 on Business Registration, Article 36.

[12] Source: HB&IS survey 2014, VASS-CAF & IRD-DIAL team.

Figure 1. Intensity of the Informal Sector by Province, 2007-2014



Source: LFSs 2007 and 2014, GSO; by courtesy of the DIAL-CAF team (Pham Minh Thai and Laure Pasquier-Doumer).

The HBs in Vietnam are highly vulnerable. Due to their micro scale, in addition to common shocks affecting all kinds of business, the HBs are also attacked by any adverse events affecting their owner or even his household members. Overall, 30 percent of HB heads experienced at least one type of detrimental events which caused significant expenditures in the last 24 months, including a severe disease (21 percent), a loss of stock or harvest to theft or climate events (seven percent), an occupational injury (four percent) or a natural disaster (two percent). Informal HBs and their owners are more vulnerable than their formal counterparts in all aspects.[13]

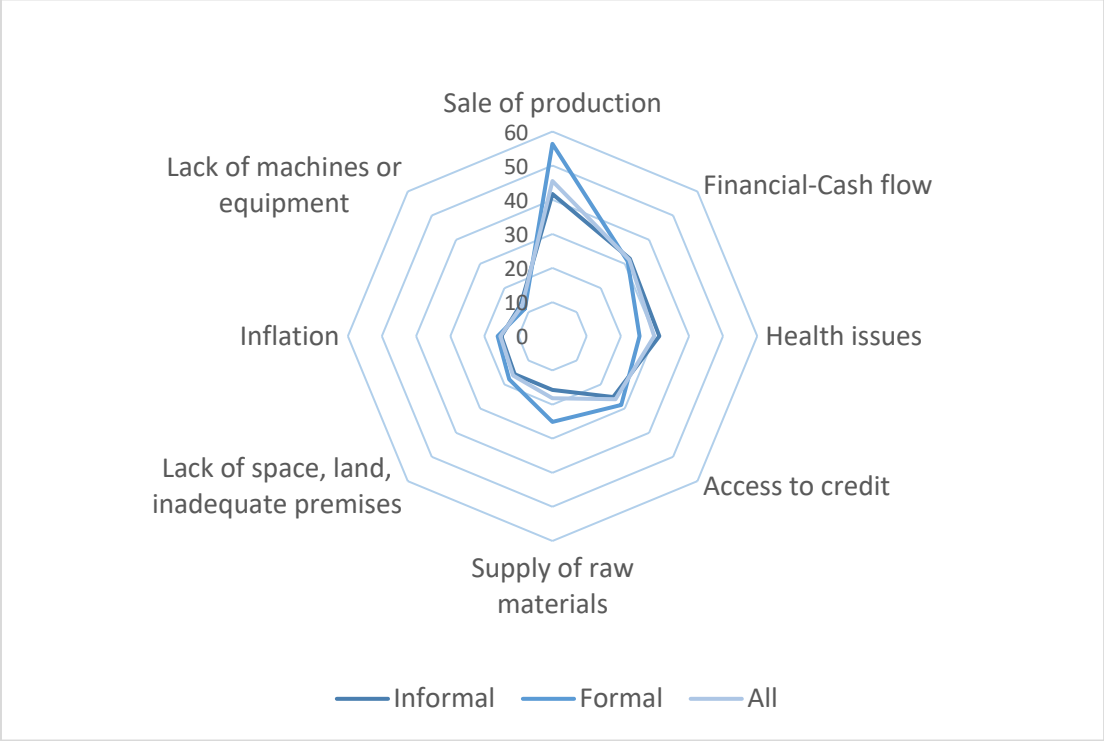
Operating a HB might confront various difficulties (**Figure 2**). Approximately eight out of ten informal HBs say that they encounter problems in doing business, and over one third of them are exposed to at least one severe problem. The most prevalent problem of the informal production units relates to the *"sale of production."* Interestingly, a greater share of formal HBs complain about the difficulty they face to sell their products and credit access than their informal counterparts, despite the fact that the former have a better economic performance and face less credit constraints than the latter.[14] On the one hand, this could be attributed to the fact that formal

[13] Source: HB&IS survey 2014, VASS-CAF & IRD-DIAL team.

[14] Source: HB&IS survey 2014, VASS-CAF & IRD-DIAL team.

HBs operate on a larger scale in a broader market, where the number of competitors is also higher (Cling et al., 2010). On the other hand, the formal production units tend to identify their business challenges better and be more ambitious than informal HBs, as a greater share of the former are planning to expand their activity in the following 12 months (37 percent) than the latter (28 percent).

Figure 2. Main Problems Encountered by HBs, by Institutional Sector (Percent of HBs)



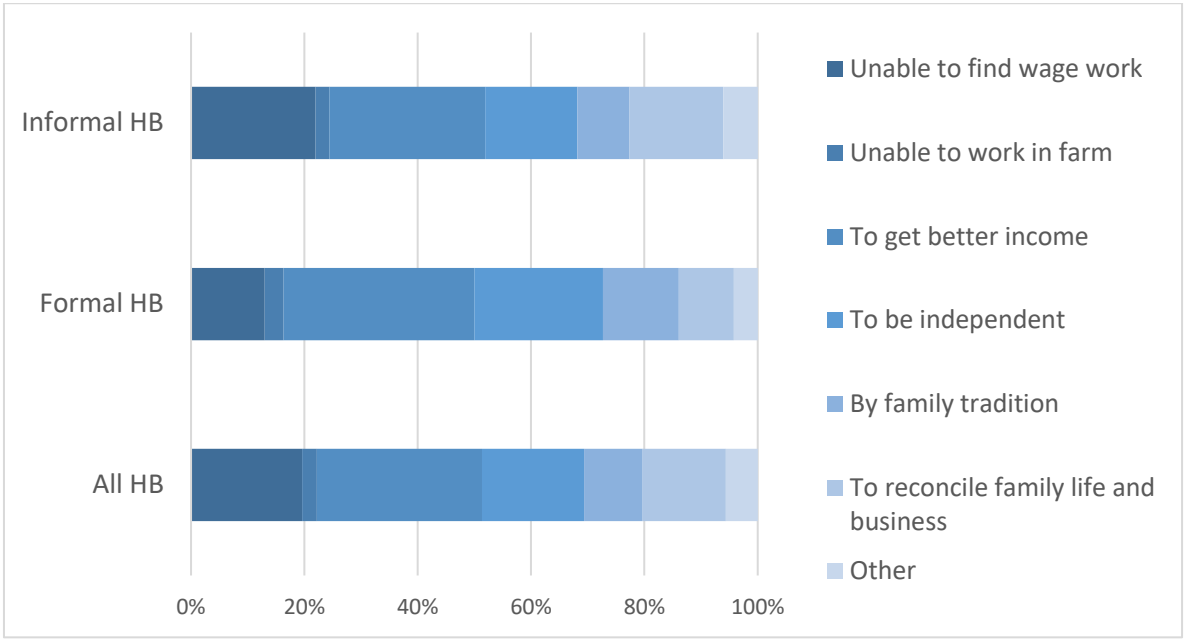
Source: HB&IS survey 2014, VASS-CAF & IRD-DIAL; authors' calculation.

Wide exposure to adverse shocks and difficulties of the HB and informal sector tend to depict a gloomy picture of this economic sphere as explained by the dualist school. According to this view, workers engage in the informal sector because they are unable to obtain a job in the formal economy. However, a number of studies have acknowledged the multifaceted nature of this sector. Fields (1990), for example, argues that the urban informal sector comprises two distinct groups, namely 'easy entry informal sector' and 'upper-tier informal sector.' The former is characterized by low wage, lack of organization and protection of workers. The latter consists of desirable activities with significant barriers to entry, higher capital or skill requirements, and fairly regular labor relation arrangements. Indeed, our statistical findings based on HB&IS survey 2014 indicate heterogeneous motivations of their involvement in the HB and informal sector (**Figure 3**).

A majority of HBs in Vietnam actually set up their business for positive reasons: to earn a better income (29 percent), to be independent (18 percent), or to reconcile family life with business activity (15 percent). Only 22 percent of HBs started their business due to a negative motivation (unable to find another job elsewhere). There are some differences between informal and formal HBs. *"To reconcile family life and business"* appears to be an important motivation of informal HBs (17 percent) rather than of formal ones (10 percent). *"Family tradition"*, by contrast, is a fairly relevant

motivation concerning formal production units (13 percent) rather than informal HBs (9 percent). Traditional craft-village HBs could possibly serve as a typical example.

Figure 3. Motivation of Setting Up an HB (percent of HB)



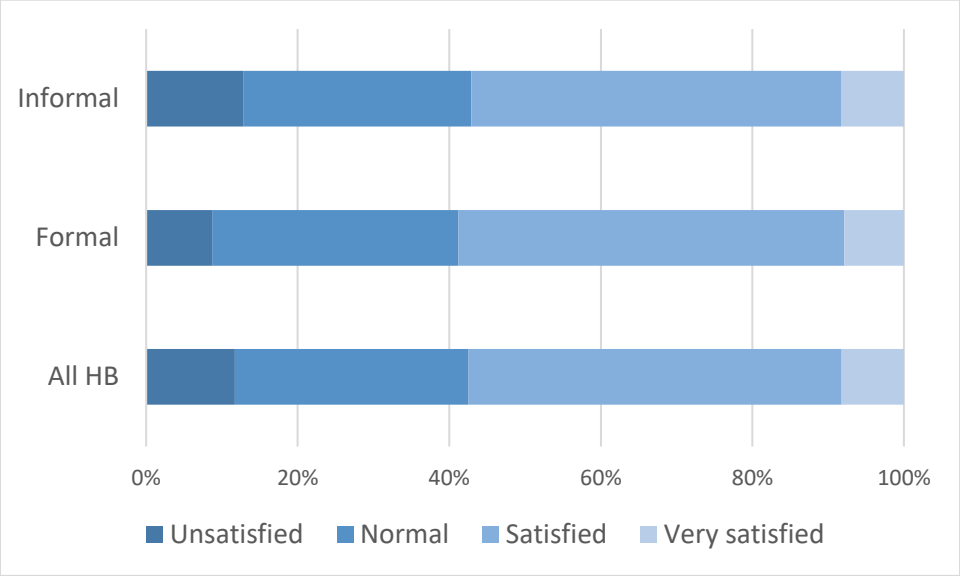
Source: HB&IS survey 2014, VASS-CAF & IRD-DIAL; author's calculation.

More surprisingly, the data show that a large share of HB heads feel satisfied working in the HB and informal sector (**Figure 4**).[15] A scale question on job satisfaction is introduced in the 2014 survey questionnaire: *"All things considered, how satisfied are you with this job?"* Five modalities are proposed: *"very unsatisfied," "unsatisfied," "normal," "satisfied"* and *"very unsatisfied."* Over a half of HBs are contented, and nearly one third of them find it normal to work in the HBs, leaving merely 12 percent of HB heads feeling unsatisfied with their business. The informal HB heads express no less satisfaction than the HB average, with 57 percent feeling satisfied or very satisfied with their business. This figure is close to that of Razafindrakoto et al. (2012) based on LFS 2009, at 54 percent.

Overall, these above results point out that despite a high vulnerability in general, the informal sector in Vietnam is not simply a residual component existing involuntarily and having inferior characteristics which should be eliminated or formalized at all prices. In fact, a significant share of HB heads find satisfied and voluntary to work in the informal sector. In addition, formal HBs declare no less difficulty than informal counterparts, and it seems that the type of problems encountered in running a HB could reveal something about their economic performance. Therefore, it would be meaningful to investigate the diversity of the informal sector, particularly starting from the issues facing their business. In this regard, the following paragraph briefly presents the first chapter which is dedicated to the heterogeneity of the informal sector in Vietnam.

[15] As the proportion of *"very unsatisfied"* responses is minimal (1.1 percent), it is combined with *"unsatisfied"* modality in our statistical analysis.

Figure 4. Satisfaction of Working in the HB and Informal Sector



Note: "Unsatisfied" category includes "unsatisfied" and "very unsatisfied" responses. Source: HB&IS survey 2014, VASS-CAF & IRD-DIAL; author' calculation.

Summary of chapter: Despite ubiquity and important economic weight, the HB and informal sector in Vietnam has been understudied and almost overlooked by policy makers. Chapter 1 investigates its diversity and the role of human capital in this sector, using a unique *quali-quanti* approach. To do so, I rely primarily on the representative textual data of the 2009 HB&IS survey. This corpus provides a direct measure of the utility of HBs in both economic and non-economic dimensions, by revealing their advantages and disadvantages when running their production unit. I apply exploratory analytical methods (i.e., correspondence and clustering analysis) to limit the pre-conceived ideas in extracting the most important information from the corpus. I 'invent' a manual coding process to fully Latinize Vietnamese language in order to make it analyzable in the Data Text Mining program. I find that the textual correspondence analysis produces meaningful factor axes which can represent the entrepreneurial skill of HBs and then be used as explanatory variables in econometric models. This leads to two main findings: First, the HB and informal sector is highly heterogeneous and can be considered as a *'continuum'* in a multi-dimensional space of their disadvantages or advantages, supplemented by other characteristics of HBs and HB heads. Second, the variation in the economic performance of HBs could be explained by their heterogeneous levels of entrepreneurial skill, which is proportionate to their scale of production. The main policy implication is that a 'one-size-fits-all' scheme would not be appropriate for this highly heterogeneous sector, and that entrepreneurial skill training courses could help improve HBs' business outcome.

2.2. Chapter 2: Returns to higher education in Vietnam: Who benefit the most?

Context: Vietnamese people have a long-lasting tradition of respect for teachers and promotion of learning. According to Confucianism, three major bonds *('tam cuong')* are: King, teacher, and parents. The intellectuals and scholars were rewarded a noble position in Vietnamese feudal society. Four main social classes, ranked from top to bottom, in Vietnam's traditional society were: scholars, peasants, artisans and merchants *("sy, nong, cong, thuong")*. Vietnam's first national university, known as Imperial Academy *('Quoc Tu Giam')* was established in 1076 to educate Vietnam's bureaucrats, nobles, royalty and other members of the elite. Confucius' teaching essentially aims at cultivating and reinforcing men's virtues rather than transferring knowledge and skills. He considers an education system that focuses on the latter as to teach 'trifling skills' (tieu xao) but not to teach human. In fact, *'Dai hoc'* (university education, 大學) is one of the four classics of Confucianism, interpreted as *'dai nhan chi hoc'* (education of the venerable sages, education to become a venerable sage).

From the late 19th century to September 1945, the French colonial regime has replaced the traditional Confucian-oriented education system by a French-Vietnamese one, which invested very little in tertiary education. The whole Indochina (i.e., today's Vietnam, Laos and Cambodia) had only three universities (of Law, Medicine-Pharmacy and Sciences) located in Hanoi with 834 students, among which 628 were Vietnamese (World Bank, 2005). Then, from 1945 to the late 1980s, the higher education system of North Vietnam adopted the Soviet model while the South Vietnam followed the old colonial French model and the newly imposed American one. The Soviet model was applied nationwide after country's reunification in 1975 (Zeleza, 2016).

A study by the World Bank (2005) reports that in the early 1980s, with the dismantlement of the importance of the collective economy, the State was no longer capable to finance education. As a result, education system confronted with a serious shortage of resources. Universities, colleges and professional secondary schools even lacked a minimum budget for maintaining their regular activities and had no autonomy. Graduates were unemployed. Lecturers had to do inappropriate jobs such as pig-raising and sewing to supplement their meager income.

The *Doimoi* Reform initiated in 1986 has caused dramatic changes in the education system, particularly at the tertiary level. All possible financial sources were to be mobilized instead of complete dependence on State's budget. Therefore, the access as well as quality of higher education have undergone great improvements. The number of universities and colleges has increased from 101 in 1987 to around 480 in 2015,[16] with the establishment of numerous non-public higher educational institutions. Over the 2004-2008 period, the gross enrolment rate at the tertiary level has increased impressively from 18.3 percent to 30.2 percent (Vu, 2012). Various flexible and diverse training programs were developed to promote labor market prospective of the graduates (World Bank, 2005). From the former Soviet model with only narrowly specialized universities, the higher education system in Vietnam nowadays comprises numerous multi-field, multi-disciplinary/comprehensive universities (Nguyen & Vu, 2015). The current higher education

[16] Source: http://www.thanhniennews.com/education-youth/private-universities-struggle-to-attract-students-in-vietnam-51756.html

system in Vietnam includes university (from 4 to 6 years, depending on the field of study), college (3 years), master (from 1 to 3 years after graduating from university, depending on the field and forms of study) and doctorate education (2 to 4 years after getting master degree).

Nevertheless, a variety of problems subsist in Vietnam's higher education system that deserve more efforts to be made. In terms of access and equity, the gaps between rural and urban areas, ethnic minorities and ethnic majority as well as the poor and the rich remain large (Vu, 2012). **Table 1** shows that the educational opportunities in Vietnam are more concentrated on the richest households. More than half of the richest quintile has attended college/university, whilst only three percent of the poorest quintile has access to higher education. In terms of quality, Vietnamese higher education institutions are characterized by poor international recognition and publication records, mismatch between teaching and research, between theories and practice, outdated methodology and content, low autonomy and accountability (Nguyen & Vu, 2015). According to Science Citation Index Expanded by Thomson Reuters, in 2007, Vietnam National University (Hanoi and HCMC) had only 52 publications in peer-review journals, compared to 3598 by the National University of Singapore and 220 by the University of the Philippines (Valley & Wilkinson, 2008). **Figure 5** reveals that Vietnam's investment in research and development (R&D) in terms of budget and manpower is remarkably low compared to other countries in the region.

Table 1. Educational Attainment of the Population Aged 20 to 24 Years by Household Wealth in Viet Nam, 2011.

Educational attainment or school attendance	Household wealth quintile				
	Q1 (Poorest 20%)	Q2	Q3	Q4	Q5 (Richest 20%)
Completed primary education	16	18	22	24	20
	(16)	(34)	(56)	(80)	(100)
Completed secondary education	9	12	22	25	33
	(9)	(21)	(42)	(67)	(100)
Has attended higher education	3	6	13	27	51
	(3)	(9)	(23)	(49)	(100)

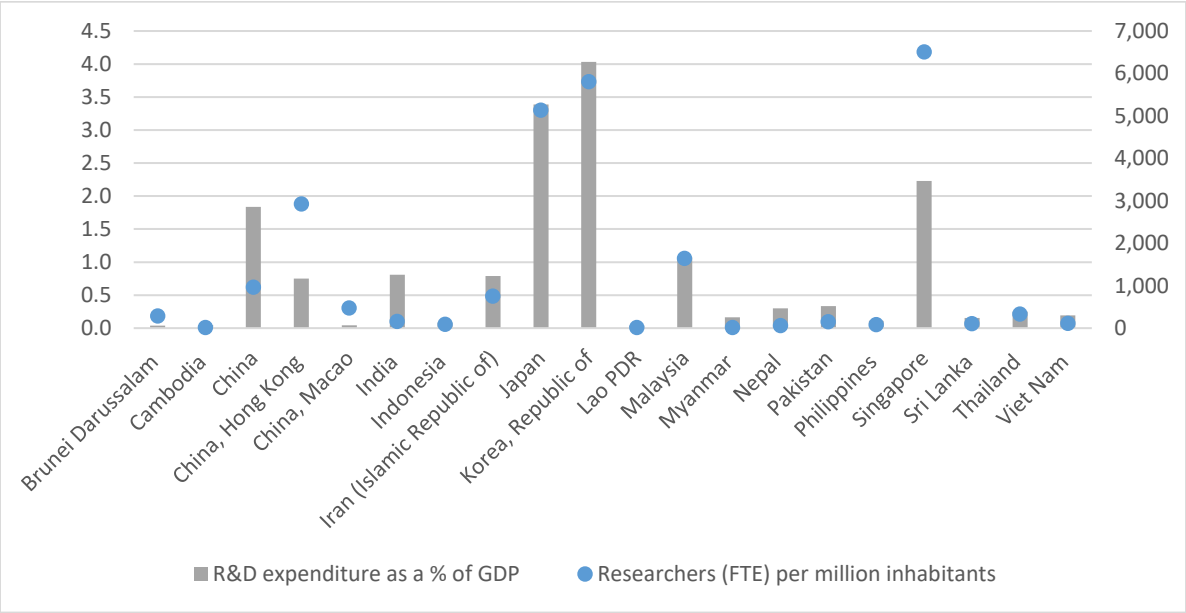
Source: UNESCO Institute for Statistics (2014). Note: Data in parentheses are cumulative percentages.

In this context, the Higher Education Law 2012 has been enacted, aiming at reforming and regulating higher education sector to develop the human resources needed for Vietnam's socioeconomic development, transformation towards a knowledge-based economy as well as international integration. The higher education reform covers all training and research disciplines and focuses on the reorganization of university and college system: renovating their structures and training scope; renewing learning and teaching, building a new system of course books, matching teaching with research, training and retraining faculty members, modernizing equipment and facilities (Nguyen & Vu, 2015).

However, these solutions seem not to touch the very essence of the current education system in Vietnam. In an essay published in Tia Sang journal, Professor Hoang Tuy[17] has acknowledged that: *"Without systematic thinking and a comprehensive, strategic vision, one could easily make himself busy with trivia and a here today there tomorrow approach, endlessly "reforming" in a fragmentary and inconsistent way, exacting huge costs but resulting in nothing more than complicating a system that is already crippled and devoid of vitality."* He regarded *"a strategic vision for immediate and long-term objectives, direction, demand, capacity relevant to development trends, guiding ideology, and a general path of actions"* as the most important factor to elevate education and science. Particularly, Professor Hoang Tuy put emphasis on moral principles or virtues of diligence, efficiency, honesty and integrity, without which an education and science system will not function properly and stagnate. Instead of implementing minor internal adjustments based on a management feedback mechanism, Professor Hoang Tuy considers to *"find and fix the system errors"* which lies in the entire state apparatus as the only way to rescue the system. According to him, eliminating the *"salary/income paradox"* is a prerequisite to ensure the four abovementioned virtues and thereby enhancing science and education standards.

In summary, the higher education in Vietnam today is characterized by remarkable improvements as well as challenges. At the country level, developing higher education is expected to bring prosperity and stability. At the individual level, attending college could be an effective means to protect one from ignorance, poverty and vulnerability. In order to make right decisions and policies, it is a must to examine the return to higher education and its variation across Vietnamese population. In this attempt, I will introduce the main findings of the second chapter in the following paragraph.

Figure 5. R&D Expenditure as a Percentage of GDP and Researchers in Full-time Equivalents (FTE) per 1 Million Inhabitants, 2011 or Most Recent Year Available



Source: UNESCO Institute for Statistics, July 2013.

[17] Professor Hoang Tuy is considered as one of the most accomplished Vietnamese scientists of the 20th century and one of the two founders of mathematics in Vietnam. He has numerous publications in international journals and a theorem bears his name. This translation is provided in Valley and Wilkinson (2008).

Summary: This chapter was motivated by the traditional, nationwide belief that a university degree is a life-changing ticket to the underprivileged. It is well known that the current higher education system in Vietnam has not met up with the requirements of the labor market. Both equality and quality concerns remain. On the one hand, facing a lack of available places, young people from disadvantaged backgrounds have less chance to be included among enrolments. On the other hand, Vietnam is suffering from an excess of low quality universities and a shortage of high quality ones, which questions the returns to higher education. It is sometimes believed that educational access and educational quality are trade-offs, for example, as public investment focuses on excellent universities (i.e., high returns) that are accessible only to the elites. To shed light on this dilemma, I investigate the variation of the returns to higher education across Vietnamese population, and essentially with individuals' probability of attending college. To do so, I rely on the Vietnam Labor Force Survey in 2007, which has a large and representative sample size. I compare the hourly earnings of those who have finished tertiary education and those who have just completed high school with comparable characteristics, using OLS, IV, Heckman, Quantile Regression and Heterogeneous Treatment Effect models. I find that on average, going to college could improve individual earnings substantially. Most importantly, contradicting to the concern about equality-efficiency tradeoff, those who are the least likely to obtain tertiary education actually tend to benefit the most from it. This result is consistent across different models. Accordingly, policies that encourage the socially disadvantaged individuals to pursue higher education may be recommended, taking into account specific circumstances.

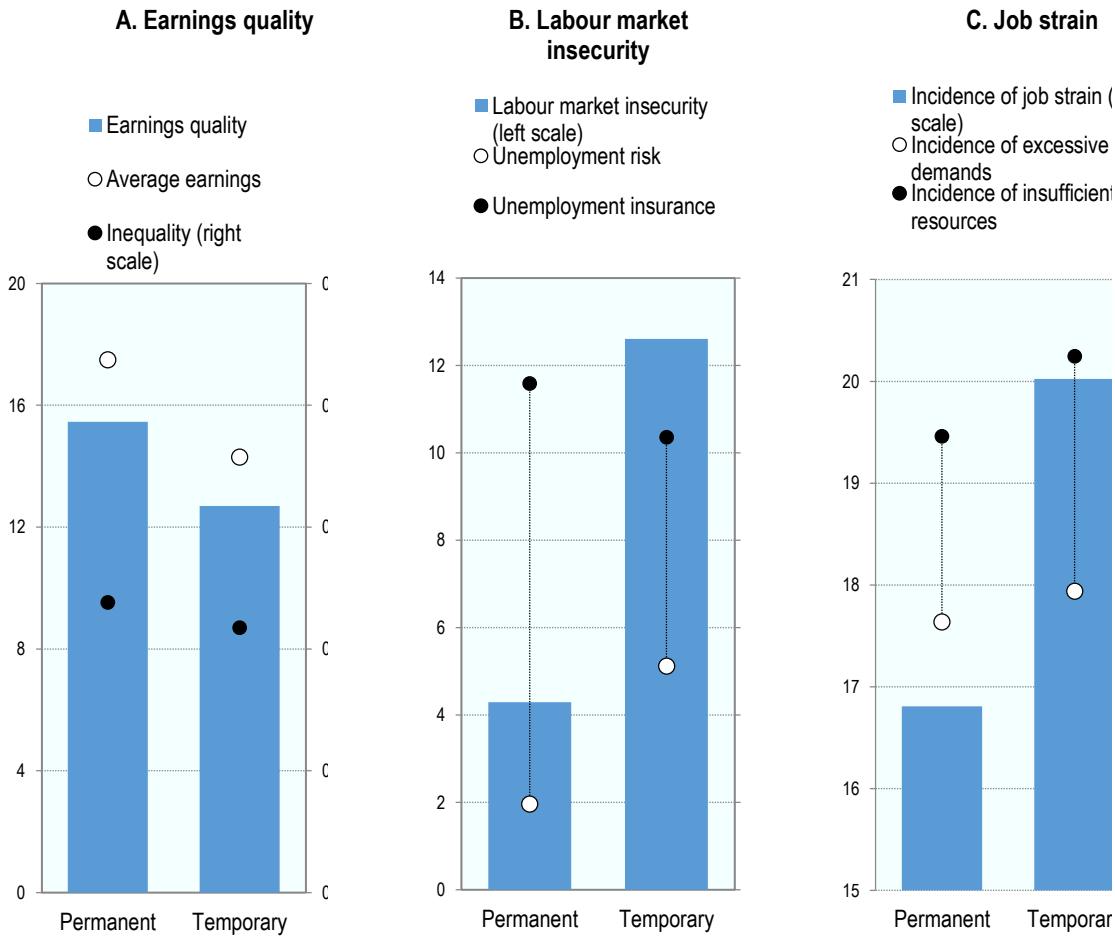
2.3. Chapter 3: Wage differentials between temporary and permanent workers in Asian developing countries

Context: The past several decades have witnessed the rise of non-standard forms of employment (NSFE), understood as jobs that do not conform to criteria of standard employment (i.e., full-time and indefinite), as well as part of a subordinate employment relationship, in many parts of both the developed and the developing worlds (ILO, 2015; Serrano et al., 2014 among others). Accordingly, NSFE includes (1) temporary employment; (2) temporary agency work and other contractual arrangements involving multiple parties, as well as (3) ambiguous employment relationships; and (4) part-time employment (ILO, 2015). Temporary employment refers to jobs with limited duration, namely fixed-term, project or task-based contracts, seasonal and casual work, including day-laborers (ibid). Although temporary employment could provide greater adjustment and flexibility, it is often associated with poor job quality, in terms of lower earnings, higher level of labor market insecurity and higher job strain (**Figure 6**).

Despite their ubiquity and potential damages to workers, accurate cross-country statistics on NSFE in general and on temporary employment in particular, as well as empirical evidence on wage differentials between temporary and permanent workers are nearly absent in developing Asian countries. World of Work Report 2008 (ILO, 2008), for example, states that: *"In emerging economies and developing countries, non-standard work mainly takes the form of informal employment, that is, workers in very small firms (fewer than five workers), self-employment, unpaid family work and salaried employment without a proper work contract in the formal sector. Unfortunately, data on informal employment are scarce and difficult to compare across countries."* A possible explanation for the lack of Asian cross-country statistics and studies on NSFE and temporary employment could be that these forms of employment have been less of a concern compared to informal employment, that there has been a lack of awareness of the phenomenon, or that there is a various range of country definitions and of measurement methods. Serrano et al. (2014), for instance, lists a variety of terms ascribed to non-standard employment in the formal sector in ASEAN countries and points out that some of them *"overlap and even bear contradictory meanings, making it difficult to use a single term in comparing the forms and trends of non-standard employment among the study countries."* For example, in Indonesia, contract workers are workers who have employment contracts with the principal company or the user company, while outsourced workers are those who have employment contracts with the outsourcing company. In the Philippines and Singapore, contract workers can also be outsourced workers or workers dispatched by manpower supplying agencies or contractors.

Regarding empirical evidence on temporary workers' wage gaps, most available studies on Asia have been for Japan and the Republic of Korea (Ahn, 2006; Lee & Lee, 2007; Diamond, 2010; Ha & Lee, 2013 among others) or on the formal-informal wage gaps in less-developed nations (Nguyen et al., 2013; Dasgupta et al., 2015 for examples). Moreover, existing evidence mainly relies on simple 'raw' wage gaps, while the 'pure' wage gaps derived from empirical wage equations are hardly provided. Furthermore, to the best of our knowledge, evidence on the distribution of wage penalties along the wage distribution is totally non-existent in Asian developing countries.

Figure 6. Job Quality and Job Quantity Outcomes by Type of Employment: Cross-country Averages, 2010.



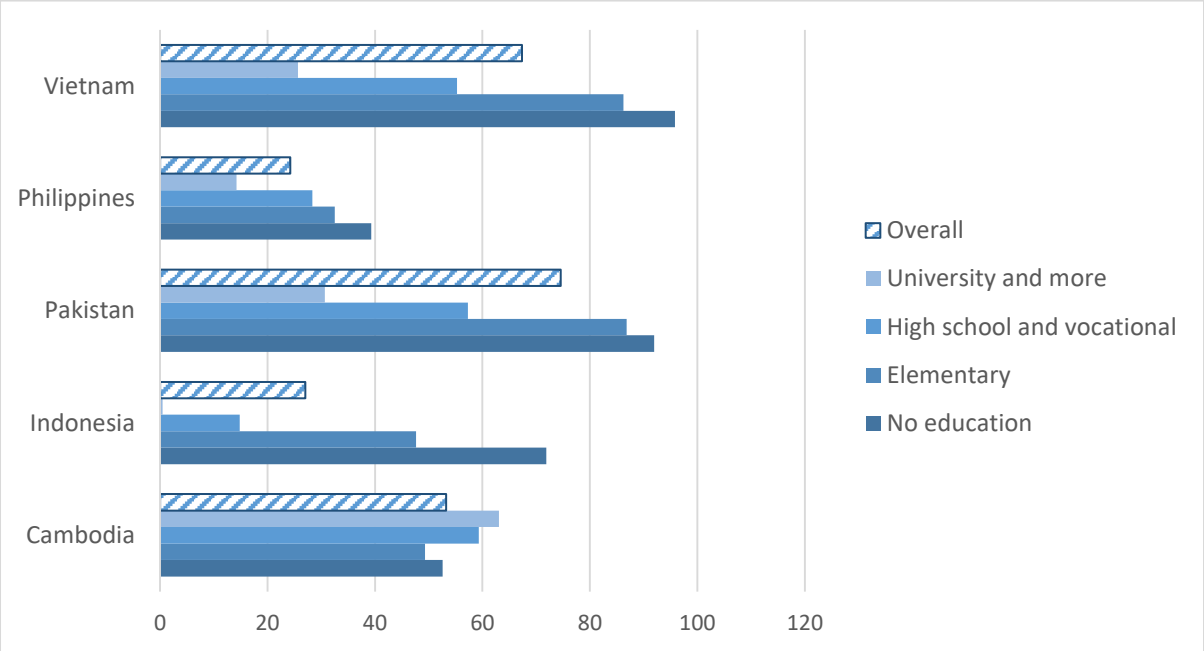
Note: Country coverage: Austria, Belgium, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Italy, Luxembourg, Netherlands, Norway, Poland, Portugal, Slovenia, Slovak Republic, Spain, Sweden, Turkey and United Kingdom.

Source: European Union Survey on Income and Living Conditions (EU-SILC); Eurofound (2012), Fifth European Working Conditions Survey, Publications Office of the European Union, Luxembourg; OECD (2014), "Labour Market Statistics: Labour force statistics by sex and age: indicators", OECD Employment and Labour Market Statistics (database), http://dx.doi.org/10.1787/data-00310-en.

In this chapter, we attempt to harmonize as much as possible the definitions of temporary employment in five Asian developing countries according to the international framework provided by the ILO. We rely on national LFS of Indonesia, Pakistan, the Philippines, Vietnam and Cambodia's Labor Force and Child Labor Survey. This chapter provides first time ever cross-country empirical evidence on the incidence of temporary employment and temporary-permanent workers' wage differentials in developing Asia. We observe that temporary jobs are prevailing in these countries, especially among the low-skilled (**Figure 7**). Evidently, education, broadly speaking human capital, is likely to affect employment choice, thus remuneration of workers as indicated in Graph 1. That is why in this study, we attempt to single out the impact of educational level and other characteristics in order to know if workers are truly discriminated just because their jobs are temporary.

Summary: In Asia, it is well acknowledged that workers could be punished financially because they are informal instead of formal, female instead of male. Yet, little is known about the remuneration discrimination against temporary workers, and most available studies on Asia have been for Japan and the Republic of Korea. Chapter 3 is the first research work that systematically examines the temporary-permanent wage gap in Asian developing countries. The analyses are based on national Labor Force Survey data of Cambodia, Indonesia, Pakistan, the Philippines and Vietnam. At mean levels, the OLS and Heckman models point out temporary employment wage penalties in all studied countries except Cambodia, where a wage premium is found. Moreover, Quantile Regression estimates show that wage differentials could greatly vary across the wage distribution. The wage gap widens at the lower tail of the wage distribution in Vietnam and Pakistan, implying that the penalty of being in temporary jobs could be more severe for disadvantaged workers, whereas the *glass ceiling* effect in Indonesia prevents temporary workers from approaching high wages. We argue that country-specific institutional and socioeconomic contexts could possibly explain the difference in the distribution pattern of wage gaps, and country-specific measures should be applied to address this issue.

Figure 7. Incidence of Temporary Workers in Selected Asian Countries, as Percentage of Wage Workers, by Level of Education



Source: Cambodia's Labor Force and Child Labor Survey 2012, Indonesia's LFS 2007-2008, Pakistan's LFS 2008-2009, the Philippines' LFS 2009, Vietnam's LFS 2007. Authors' calculation.

References

[NTDonChina]. (2012, December 21). Discovering China – The Tang Dynasty and Military Strategist Sun Bin. [Video File]. Retrieved from: https://www.youtube.com/watch?v=0fCh2UrpyXU&list=PLA6C424A04BE6060B&index=14.

Ahn, J. (2006). Nonstandard work in Japan and Korea – the origin of wage differentials. Non-standard Work-Draft. Retrieved from: http://www.jil.go.jp/profile/documents/Ahn.pdf.

Boothroyd, P. & Phạm, X. (2000). *Socioeconomic renovation in Viet Nam* (1st ed.). Ottawa: International Development Research Centre.

Chandy, L., & Gertz, G. (2011). *Poverty in numbers: The changing state of global poverty from 2005 to 2015*. Washington DC: Brookings Institution.

Cling, J., Nguyen, T., Nguyen, H., Phan, T., Razafindrakoto, M., & Roubaud, F. (2010). *The informal sector in Vietnam. A focus on Hanoi and Ho Chi Minh City*. Hanoi: The Gioi.

Dasgupta, S., Bhula-or, R., & Fakthong, T. (2015). Earnings differentials between formal and informal employment in Thailand. *ILO Asia-Pacific Working Paper Series*.

Debrah, Y. A., & Smith, I. G. (2002). *Globalization, employment and the workplace: diverse impacts* (1st ed.). London: Routledge.

Diamond, J. (2010). Non-standard Employment and Earnings in Japan. Retrieved from http://www.econ.hit-u.ac.jp/~diamond/Jess_Diamonds_Homepage/Research_files/ns%20wages.pdf

Felipe, J., & Hasan, R. (2006). *The Challenge of Job Creation in Asia*. *ERD Policy Brief Series* No. 44. Metro Manila: ADB.

Fields, G. S. (1990). Labor Market Modeling and the Urban Informal Sector: Theory and Evidence [Electronic version]. In D. Turnham, B. Salomé, & A. Schwarz (Eds.), *The Informal Sector Revisited* (pp. 49-69). Paris: Organization for Economic Co-operation and Development.

Gill, I., & Kharas, H. (2007). *An East Asian Renaissance: Ideas for Economic Growth*. Washington, DC: World Bank.

Glewwe, P., Agrawal, N., & Dollar, D. (2004). *Economic Growth, Poverty, and Household Welfare in Vietnam*. Washington, DC: World Bank.

Ha, B.-J., & Lee, S. (2013). Dual dimensions of non-regular work and SMEs in the Republic of Korea: Country case study on labor market segmentation. *ILO Employment Working Paper No. 148*.

Hart, K. (1973). Informal income opportunities and urban employment in Ghana. *Journal of Modern African Studies*, *11*(01), 61-89.

ILO (2005). *Making decent work an Asian goal.* Fourteenth Asian Regional Meeting. Busan, Republic of Korea, October 2005. Report of the Director-General. Volume II.

ILO (2008). *World of Work Report 2008: Income Inequalities in the Age of Financial Globalization, International Institute for Labour Studies.* Geneva: International Institute for Labour Studies.

ILO (2012). *Global Employment Trends 2012*. Geneva: ILO.

ILO (2015). *Non-standard forms of employment.* Geneva: ILO.

Kohli, H. S., Sharma, A., & Sood, A. (Eds.). (2011). *Asia 2050: realizing the Asian century*. New Delhi: SAGE Publications.

Lee, B.-H. & Lee, S. (2007). Minding the gaps: Non-regular employment and labour market segmentation in the Republic of Korea. *ILO Conditions of Work and Employment Series 19*.

Nguyen, H., Nordman, C., & Roubaud, F. (2013). Who Suffers the Penalty?: A Panel Data Analysis of Earnings Gaps in Vietnam. *Journal Of Development Studies*, *49*(12), 1694-1710.

Nguyen, V. N., & Vu, N. T. (2015). Higher Education Reform in Vietnam: Current Situation, Challenges and Solutions. *VNU Journal of Science: Social Sciences and Humanities*, *31*(4), 85-97.

OECD (2014). *OECD Employment Outlook 2014*. OECD Publishing. http://dx.doi.org/10.1787/empl_outlook-2014-en

Ofreneo, R. E. (2013). *Asia and the Pacific: Advancing Decent Work Amidst Deepening Inequalities*. [Singapore]: ITUC-Asia Pacific.

Razafindrakoto, M., Roubaud, F., & Wachsberger, J.-M. (2012), "Satisfaction dans l'emploi et le secteur informel au Vietnam", in: J.P. Cling, S. Lagrée, M. Razafindrakoto and F. Roubaud (eds). *L'économie informelle dans les pays en développement*, Collections Conférences and séminaires, N°6, Agence Française de Développement, Paris.

Serrano, M., Marasigan, M., & Pupos, V. (2014). *Between Flexibility and Security: The Rise of Non-Standard Employment in Selected ASEAN Countries*. Jakarta: ASEAN Services Employees Trade Unions Council (ASETUC).

Standing, G. (2011). *The precariat: The new dangerous class*. London and New York: Bloomsbury Academic.

UN Habitat/ UNESCAP (2010). State of Asian Cities Report 2010/11. Fukuoka: UN-Habitat.

Vallely, T. J., & Wilkinson, B. (2008). Vietnamese higher education: Crisis and response. *Memorandum: Higher Education Task Force, Harvard Kennedy School*. Retrieved from: https://www.hks.harvard.edu/innovations/asia/Documents/HigherEducationOverview112008.pdf.

Vu, H. L. (2012). An overview of access to and inequality in the education system of Viet Nam. *Asia-Pacific Development Journal*, *19*(1), 37-62. http://dx.doi.org/10.18356/57405f5e-en.

World Bank (1998). Vietnam Development Report 1999: Preparing for Take-offs? Hanoi.

World Bank (2005). *Education in Vietnam: Development history, challenges and solutions*. Retrieved from: http://siteresources.worldbank.org/EDUCATION/Resources/278200-1121703274255/1439264-1153425508901/Education_Vietnam_Development.pdf

Zeleza, P. T. (2016). *The Transformation of Global Higher Education, 1945-2015*. New York: Palgrave Macmillan.

Chapter 1

Heterogeneity of the Informal Sector in Vietnam: A Quali-quanti Approach

Abstract

The knowledge in the informal sector in Vietnam is paradoxically limited despite its economic weight. This paper sheds light on the heterogeneity of this sector, using a unique *quali-quanti* approach. The contributions of this paper are threefold: (1) The representative textual data provide a direct measure of utility of household businesses in both economic and non-economic dimensions; (2) This is the very first economic paper that solves the linguistic problems associated with Vietnamese textual data; and (3) Exploratory analytical methods used in this paper (i.e., correspondence and clustering analysis) help limit the pre-conceived ideas of researchers. The textual data are 'quantified' into principal axes, which are then used as continuous variables to explain the functioning and economic achievement of HBs in econometric models. We find that HBs and informal sector is highly heterogeneous and can be considered as a *'continuum'* in a multi-dimensional space of their disadvantages or advantages. In addition, the variation in the economic performance of HBs could be explained by their heterogeneous levels of entrepreneurial capital, which is proportionate to their scale of production. The main policy implication is that a 'one-size-fits-all' scheme would not be appropriate for this highly heterogeneous sector, and that entrepreneurial skill training courses could help improve HBs' business outcome.

Acknowledgements

This chapter owes much to the Vietnamese-French team who participated in the project on the Household Business and Informal Sector in Vietnam and its sponsors. Useful comments were received at the ARCUS conference on "Population and Development in Vietnam: Recent results" in Paris, France; at the 7th Vietnam Economist Annual Meeting (VEAM), HCMC, Vietnam; and at the 21st Economic and Policy Research Seminar of the Vietnam Center for Economic and Policy Research (VEPR) in Hanoi, Vietnam. A related work of this paper, co-authored with Mireille Razafindrakoto and François Roubaud, has been published as a chapter in the book "Mutations démographiques et sociales du Vietnam contemporain" by the University of Paris Nanterre in 2016; for which we would like to thank the editors and an anonymous reviewer.

1. Introduction

The informal sector is ubiquitous in Vietnam as in most developing countries. In 2014, with 6.5 million production units, the informal sector is the second job supplier after agriculture sector, accounting for nearly 40 percent of non-farm employment in Vietnam.[18] Paradoxically, this sector attracts too little attention from the State as well as from researchers. On the State's side, none of the temporary measures included in the stimulus package to abate the negative impact of the global crisis has benefitted the informal sector (Cling et al., 2010a). On the research side, few of the existing literature accurately study the informal sector according to its international definition. Most of them are based on ad-hoc partial surveys that only cover a few hundred businesses that concentrated in certain activities and places (Cling et al., 2012). This paper contributes to the knowledge in this area by investigating the heterogeneity of this sector, based on an accurate, representative and unique set of qualitative and quantitative data.

Among literature on informality, there are two important questions that have been broached to debate. The first question focuses on whether the informal sector is homogeneous or heterogeneous; and how heterogeneous it is. The second one addresses the voluntariness of engagement into the informal economy, i.e., whether it is a choice or a constraint. This paper aims at the first question, and to some extent, draws the link with the second one. By following a *quali-quanti* approach, correspondence and cluster analyses are conducted on textual responses to the open-ended questions, then the empirical results are enriched with categorical variables. The main finding is that the informal sector can be considered as a *continuum* of household businesses (HBs) in the multi-dimensional space. It is far from homogeneous, with different groups of HBs encountering different constraints and advantages as well as having different characteristics. It also appears to be more complex than dual segmentation view. We also find that the variation in the economic performance of HBs could be explained by their heterogeneous entrepreneurial skill, measured by principal factors derived from correspondence analysis.

The rest of the paper is structured as follows. Section 2 provides an exhaustive review of the literature on informality with an emphasis on Vietnam. It also shows the novelty of this research work in comparison with existing studies. Section 3 presents data and the methodology adopted. In particular, it explains how the input textual data are generated by a new coding method that the author developed, while taking into account the specific challenges of Vietnamese language. Empirical findings are reported in Section 4, including descriptive statistics, correspondence analysis and cluster analysis with a rich visualization, econometric analysis as well as robustness checks. Section 5 delivers policy implication and conclusion.

[18] Source: HB&IS survey 2014, VASS-CAF & IRD-DIAL team.

2. Literature review

Existing literature on informality can be divided into two broad categories: the homogenous and the heterogeneous views. In both categories, the complexity of informality has been so far acknowledged,[19] but at different levels of emphasis. The first category, to some extent, considers the informal sector as a single entity and focuses on its distinction from the formal sector; whereas, the latter insists more on the multi-facet nature of the informal sector, i.e., points out the differences between various segments within this complex sector.

The homogeneous view appears to give a gloomy picture of the informal sector. This sector is characterized by generally small and unproductive firms, which are run by less educated managers, use less capital and finance, have different customers from those of formal firms and rarely become formal (La Porta & Shliefer, 2008). There is overwhelming evidence that informal sector players have limited or lack of access to resources and market as well as to land and physical infrastructure (Sethuraman, 1997; Tokman, 1990, as cited in Blunch et al., 2001). In Vietnam, the informal sector consists mostly of self-employed people having very low income. It systematically employs the least educated urban population and strikingly, is narrowly integrated into the rest of the economy (Cling et al., 2010a). Workers in the informal sector are characterized by a low level of job satisfaction (Razafindrakoto et al., 2012). Competition is the main problem in this sector (Cling et al., 2010a). The origins and motivations of informality have triggered controversy among three dominant schools of thoughts (Bacchetta et al., 2009). The *'dualist'* approach is based on a dual labor market model proposed by Lewis (1954) and Harris and Todaro (1970). This model considers the informal sector as a residual component, which only exists to absorb the unemployed by the formal economy. Meanwhile, the *'structuralist'* acknowledges the interdependencies between the informal and formal sectors (Moser, 1978; Portes, Castells & Benton, 1989). The former provides the latter with cheap labor and products, thereby improving their competitiveness. The *'legalist*,' in contrast to the other two, emphasizes the voluntariness of firms to be informal (de Soto, 1989; Maloney, 2004). Under this approach, micro-enterprises choose to operate informally to evade regulation burden.

Over the past two decades, the debate on whether informality is a constraint or a choice has increasingly been polarized. Vietnam is not an exception. On the one hand, favorers of dualism argue that HBs make up a low productivity sector that serves as a safety net for the poor (Oostendorp et al., 2009) or registration is causally beneficial to both informal firms and workers (Rand & Torm, 2012). On the other hand, many studies support the legalism that cumbersome public regulations hinder the formalization process (Taussig & Pham, 2004; Nguyen & Pham, 2006).

The heterogeneous view: Seeking to reconcile the aforementioned contrasting views, the second category of literature presents the informal sector as a multi-segmented phenomenon (Bacchetta et al., 2009). Fields (1990), for example, argues that urban informal sector consists of

[19] Maloney (2004), for example, despite arguing that being in the informal sector is a voluntary choice, stated that, *"…it is critical to acknowledge that the informal sector is very heterogeneous and the view I am laying out here will not describe well the peculiarities of every subsector."*

two distinct groups, namely 'easy entry informal sector' and 'upper-tier informal sector.' The former is characterized by low wage, lack of organization and protection of workers. People in this sub-sector seek to get out of it. The latter provides desirable activities with significant barriers to entry, higher capital or skill requirements, and fairly regular labor relation arrangements. Recent empirical evidence supports his argument. In a study on Brazil, Mexico and South Africa, Bargain and Kwenda (2011) point out the dual nature of the informal sector, with self-employed workers in the high-tier segment receive a significant earnings premium, while workers at the lower end of the earnings distribution face significant earnings penalties compared to the formal sector. In another study, Cunningham and Maloney (2001) classify Mexican micro firms into six distinct clusters by using factor and cluster analyses. Recent studies on Egypt (Harati, 2013), Turkey (Aydın et al., 2010) also suggest the existence of different segments inside the informal sector.

Particularly, Benjamin and Mbaye (2012) raise a provocative conception, wherein the informal sector is considered as a *continuum*. At the bottom of the scale, purely informal companies do not meet any criterion of formality. The second level consists of firms that meet at least one of the criteria defining formality and so on. The highest-level firms are completely formal. In this paper, the concept of *continuum* is not associated with the level of formality since business registration is the single clear cut between informal and formal HBs. Instead, this continuum is fabricated from HB heads' responses to the open-ended questions which reveal their advantages and disadvantages in running their business.

In Vietnam, recent research based on representative data on the informal sector seems to be in favor of the heterogeneous view. Cling et al. (2014) propose a 'pyramid' of the labor market that comprises five main segments. The informal sector is broken into 'upper-tier' and 'free-entry' categories. In this pyramid, the wage workers in formal enterprises and HBs are squeezed between these two informal sub-sectors, suggesting that the income level of upper-tier informal HBs is even superior to that of formal wage employees.

In particular, Cling et al. (2012) show that there is a difference between businesses which are set up to be independent or to follow a family tradition, and those set up for a lack of an alternative or to make extra income for the household (i.e., auxiliary activity). The former have a higher probability of being formal. In another research work, Cling et al. (2010b) apply multiple correspondence and clustering analyses on the quantitative data of numerous HBs' characteristics to distinguish three main groups of HBs at different levels of insecurity and professionalism: *'the survivor,'* which are the most insecure and precarious; *'the resourceful,'* and *'the professional,'* which are the high-end group. In addition, based on the quantitative data of problems that HBs encounter, they identify four clusters of HBs, namely: (i) *'The ambitious,'* who express dissatisfaction with their equipment and plan to enlarge their business; (ii) *'The optimistic,'* who aspire an improved business environment (regulation and access to services) and believe there is a future for their activity; (iii) *'The pragmatic,'* which state mainly competition and market access problems, but seem more accommodating and do not really require assistance; and finally (iv) *'The fatalistic,'* who see no reason to complain and do not expect any change. This four-category classification implies a continuum of expectation of HBs, which appears to be correlated with their level of economic performance.

Nonetheless, there are two key limits in these analyses. Firstly, they are based on indirect measures of utility of HBs. The researchers made an implicit assumption that HB's utility essentially depends upon their economic and business outcomes. Their utility, however, may depend on other non-monetary aspects that can hardly be captured in quantitative data. Secondly, the researchers imposed their *à priori* conception of the nature of HBs in this sector. That is, for example, to pre-determine the list of difficulties faced by HBs in a close-ended question. This paper overcomes the shortcomings of the previous literature by exploiting three open-ended questions regarding the advantages and disadvantages of HBs. These questions, on the one hand, directly reveal the utility of HBs on multiple dimensions (both economic and non-economic). On the other hand, they give HBs full freedom to express their experience, without pre-conceived ideas of the researchers involved.

3. Data and methodology

3.1. Data

This research is based on the Household Businesses and Informal Sector Survey (HB&IS) conducted in 2009 in Ho Chi Minh City (HCMC), the number-one economic hub in Vietnam.[20] The HB&IS Survey adopts an appropriate definition of the informal sector in Vietnam, which is in line with international recommendations, and covers a representative sample of HBs. The originality of the data is marked by three open-ended questions, which are, to my knowledge, added into an informal sector survey for the very first time. In other studies on the informal sector, qualitative information has been collected, but mainly for a subgroup of key informants (e.g., expert interview) or case studies.

To start with, the informal sector is defined as all private unincorporated enterprises that: (i) produce at least some of their goods and services for sale or barter; (ii) are not registered (i.e., having no business license); and (iii) are engaged in non-agricultural activities (Cling et al., 2010b). In Vietnam, 'unincorporated enterprises' are called 'household businesses,' which either have to be registered (i.e., formal) or not (i.e., informal) depending on their income level (Box 1).

The HB&IS Survey is the second phase of a *1-2-3* scheme. *1-2-3* is a modular mixed survey, whose general principle is to use the information drawn from a household survey concerning individuals' activity (Phase 1: Labor Force Survey) to select a sample of production units (PUs) to which is applied a specific questionnaire on informal activities (Phase 2: HB&IS Survey). The basic tenet of mixed informal sector surveys is that informal PUs are easier to capture through the jobs of the people working in the informal sector than via the identification of the premises on which the businesses are run as in Establishment surveys (Cling et al., 2010b). In Vietnam, the survey scope was extended to all HBs to allow for a comparison between informal and formal PUs. Section 4, after presenting the findings for informal HBs, will expose a broader picture of all HBs.

[20] HB&IS Survey was jointly conducted by General Statistics Office (GSO Vietnam) and Institut de Recherche pour le Développement ; Développement, Institutions et Mondialisation (IRD-DIAL France).

In 2009, there were actually two different HB&IS surveys in HCMC: one on the new random sample drawn from LFS 2009, and the other on the existing random sample drawn from LFS 2007. The latter survey is thus considered the second wave of HB&IS panel data. This survey primarily covers 1333 HBs, which were randomly chosen and first interviewed in 2007. However, only 1023 HBs of them sustained in 2009, among which 699 HBs were informal and 324 HBs were formal. Attrition from 2007 to 2009 is non-random. According to Cling et al. (2010a), those who closed their business between 2007 and 2009 were more likely to operate on the streets (i.e., more precarious) than on average. HBs operating in the services sector have the highest rate of closure and in manufacturing the lowest. Meanwhile, the size of HBs or the age of their heads does not seem to influence their probability of closure.

Box 1. Business Registration of Household Businesses in Vietnam

(Decree N°88 dated 29/08/2006 on Business Registration, Article 36)

1. A household business as owned by one Vietnamese citizen, by one group of persons or one individual household may be registered for business at one location only, may employ only up to ten employees, shall not have a seal, and shall be liable for its business activities to the full extent of its assets.

2. Household businesses which engage in agricultural, forestry, fishery or salt production or which are street hawkers, nosh vendors, long-distance traders, itinerant traders or service providers earning low income shall not be required to register their businesses, unless they conduct business in conditional lines of business. People's committees of cities and provinces under central authority shall stipulate the applicable level of low income within their locality. The stipulated level of low income may not be higher than the stipulated threshold at which personal income tax is payable in accordance with the law on tax.

3. Any household business which employs regularly more than ten employees has to register business as an enterprise.



Source: Economica Vietnam (2013)

There is also a separate qualitative survey in Hanoi and HCMC in 2009 and 2010 aiming at complement the quantitative information brought by the LFS and HB&IS in 2007 and 2009. 60 HBs drawn from the HB&IS panel sample were interviewed as case studies. This survey was to further investigate the characteristics of the informal HBs, their motivations of engaging in this sector and their functioning, as well as the impacts of the economic crisis on the informal sector (Razafindrakoto & Nguyen, 2010).

We use the 2009 wave of HB&IS panel survey for this study because this is the only dataset with the original open-ended questions on a large number of HBs. Therefore, it must be kept in mind that the obtained results are only representative for the HBs which existed in 2007 and survived in 2009. Moreover, to provide a vivid description of the informal sector, the detailed responses to the qualitative survey conducted on 60 HBs in Hanoi and HCMC in 2009 and 2010 are also mobilized (Section 4.1).

Quantitative data: are derived from five modules, namely: (i) characteristics of the establishments; (ii) labor force; (iii) capital, investment and financing; (iv) customers, suppliers and competitors; and (v) problems and prospects. Continuous variables are transformed into categories, and some modalities of categorical variables are regrouped so that they are more relevant to the analyses. The variables created and used in the analyses are listed in **Table 18** (Appendix) and constitute two groups: Characteristics of HBs and of HB heads.

Qualitative data: comprise three open-ended questions as follows:

(1) Please state any advantages while running a business like yours.

(2) Please state any disadvantages while running a business like yours.

(3) State any subjects related to your business that have not been mentioned or clarified before.

The textual data provide unique, copious and intrinsic information on the characteristics and utility of the HBs, by giving room for them to express their own experience. These kinds of information are rarely acquired in the quantitative surveys. Yet, the texts require a careful cleaning procedure due to some issues. For examples, spelling errors usually occur; a number of sentences and phrases are left incomplete; some information is misplaced (e.g., a disadvantage is reported in 'advantage' variable) and so on. Moreover, very often, the information collected in the third question explicitly reveals their disadvantages and advantages. Therefore, it was a must to 'clean' the responses one-by-one, although this work was time-consuming. In order to minimize our subjective influence to the data, the correction is confined within the followings:

(1) Remove the responses of HBs who had closed their business.[21]

(2) Correct spelling mistakes.

[21] This is because: Firstly, HBs that were closed did not provide any other information except to the qualitative survey. Secondly, the number of closed HBs who answered the open-ended questions is very limited. In summary, including these HBs into the analysis adds little information on the one hand, and biases the scope of the analysis (i.e., HBs which existed in 2007 and survived in 2009) on the other hand.

(3) Move the advantages mistakenly reported in 'disadvantage' variable to the 'advantage' variable and vice versa.

(4) Add explicit disadvantages and advantages revealed by the third question to the first two variables.

After this cleaning step, the texts have to be codified such that the Text Mining software program can work on them. This is a challenge specific to Vietnamese language, which will be presented in the next section.

3.2. Methodology

This paper employs a combination of quantitative and qualitative approaches (i.e., so-called *quali-quanti* approach). The quantitative approach is appreciated for its precision and ability to generalize, while being criticized for its restricted freedom and imposed *à priori* conceptions of researchers. Meanwhile, the qualitative approach often raises the problem of subjectivity and lack of generalizability. However, it gives the respondents the freedom to express their own experience, thus suggesting new insights (Hughes, 2012). When used together, these methods can be complimentary.[22] In this research, the interaction between these two approaches is reflected in both data collection and data analysis. First, a quantitative approach (i.e., questionnaire, random sampling) collects qualitative data through open-ended questions along with quantitative data. Second, qualitative data are quantified by a quantitative analytical tool (i.e., DTM-Vic software). Also, the correspondence and cluster analysis on texts is supplemented by categorical variables. Moreover, the findings from one type of study are used to check against the findings derived from the other type (i.e., 'logic of triangulation' - Bryman, 2006).

The analyses are performed mainly in DTM-Vic version 5.6, a program developed by L. Lebart et al. (2012), which is devoted to the exploratory analysis of multi-dimensional data. The basic tools that are employed in this paper are correspondence and clustering analysis.

Correspondence analysis is conceptually similar to Principal Component Analysis (PCA). It involves a mathematical procedure that transforms a number of possibly correlated variables into a smaller number of uncorrelated variables called principal components. The components are created to account for maximal variation among the original variables, i.e., the first principal component accounts for as much of the variability in the original variables as possible, and each succeeding component accounts for as much of the remaining variability as possible (Lebart et al., 2006).

Clustering is a division of data into groups of similar objects. Clusters are formed such that HBs in the same cluster are as similar to each other as possible, and that HBs in different clusters are as distinct as they can be (Wanner, 2004). Clustering can be performed on the principal components of PCA "to denoise the data" (Husson et al., 2010). In DTM-Vic version 5.6,

[22] For a typology of *quali-quanti* approaches, see more in Teddlie & Tashakkori (2008) and Bryman (2006).

clustering is based on a hybrid method using both bottom-up hierarchical (agglomerative) and k-means clustering (Lebart et al., 2006).[23]

According to Lebart and Mirkin (1993), correspondence analysis and clustering are practically complements and it seems wiser to use both. They argue that correspondence analysis could result in shrinkages and distortions due to both the projection onto the principal dimensions and the possible lack of robustness of the global fit (sensitivity to outliers). It is thus desirable to complement it with a classification performed in the whole space. Being derived in a much higher dimensional space allows them to provide the information that could have been obscured by the projection onto a low dimensional subspace. Moreover, it is much easier to describe a set of clusters than a continuous space.

Interestingly, these techniques can be applied to textual analysis. The objective is to describe, explore and extract the main information contained in the texts. The text mining techniques have been applied successfully in the analyses of speeches, political opinions and ancient texts. This extension requires a transformation of texts into numerical data. This step cross-tabulates a lexical table with columns being distinct words used and rows being HBs (i.e., 'contingency table'). The cell value is thus the frequency with which a certain word appears in a response (Lebart et al., 1998; Yelland, 2010). Correspondence and clustering analyses are then implemented just as on categorical data.

Vietnamese language – a challenge

Economic research employing these text mining techniques seems to be rare. It is even non-existent for Vietnamese texts. Until now, statistical programs such as DTM-Vic have not been able to work properly with Vietnamese. This is due to specific features of this language as explained below.

(i) Alphabet

The Vietnamese alphabet contains special letters that are non-existent in the classical Latin alphabet such as ê, ă, â, ô, ơ, ư.

(ii) Tone marks

Vietnamese has six tones: *level, hanging, sharp, asking, tumbling,* and *heavy*. They are expressed by diacritic marks. Words with similar letters may have totally different meanings while marked by different tone marks. For example, 'tôi' (I, myself) changes its semantic content to be 'bad' (tồi), 'dark' (tối) or 'crime' (tội). Therefore, ignoring those marks in the analyses is impossible.

[23] Agglomerative clustering algorithm starts with one-point (singleton) clusters and recursively merges two most similar clusters. The similarity matrix is then updated to reflect the pairwise similarity between the new cluster and the remaining clusters. These steps are repeated until a single cluster remains. The output of this procedure is a hierarchical tree or *"dendrogram."* In K-means clustering method, however, k data points are randomly chosen as the initial centroids. All points are then assigned to their closest centroids and centroid of each newly assembled cluster is recomputed. This process is repeated until the centroids do not change.

(iii) Word boundary

Unlike English or French, where *"Word is a single distinct meaningful element of speech or writing, used with others (or sometimes alone) to form a sentence and typically shown with a space on either side when written or printed"* ("Word," 2013), in Vietnamese, whitespaces are not used to identify the word boundaries. Indeed, Vietnamese is monosyllabic in nature. Each 'syllable' (i.e., 'tiếng') tends to have its own meaning and is written with a space before and after. However, Vietnamese words are at times made of two or more syllables. Thus, 'morphosyllables' in Vietnamese are considered to have the status of 'morphemes' (Tran et al., 2007). Words made from two or more morphemes can only be recognized by context. DTM-Vic may consider morphemes as words, thus misunderstanding their significations. For example, in the sentence: "Gia đình tôi gặp khó khăn về tài chính" (My family faces financial difficulties), 'tài' could imply 'talent,' 'chính' possibly means 'main,' but a combination of 'tài chính' turns out to be 'finance.' In this context, it must be the whole 'tài chính' as a single word to be correctly meaningful.

New coding method

Cutting-edge research in linguistics and informatics has developed various methods to tackle the issue of word segmentation in Vietnamese, such as Dinh et al. (2001); Nguyen, C. T. et al. (2006) and Nguyen, T. V. et al. (2006). Yet, none of them are perfect. Moreover, the first two problems remain in DTM-Vic. Taking into account the sample size and the length of responses, I decided to code the texts manually in a way that addresses the aforementioned problems:

(i) The alphabet problem (i.e., 'chữ cái') was solved by replacing special letters by a pair of Latin letters that imply no other meaning in Vietnamese:

ă = aw; â = aa; đ = dd; ê = ee; ô = oo; ơ = ow; ư = w

(ii) The tone marks problem ('dấu') was tackled by the same logic: placing a specific Latin letter, which never ends a word in Vietnamese, for each tone:

` (*hanging*) = f, ´ (*sharp*) = s, ˀ (*asking*) = r , ~ (*tumbling*) = x, . (*heavy*) = j

The idea to deal with the two first problems comes from the Vietnamese typing rule with English keyboard of Vietkey and Unikey software programs.

(iii) Separate morphemes of a compound word are identified and linked with each other, making compound words continuous strings of pure Latin letters.

E.g., 'cạnh tranh' = canhjtranh (competition), 'rảnh rỗi' = ranhrroix (free, idle)

After the texts are coded to be usable in DTM-Vic, they need to be further refined before applying the correspondence and clustering analyses. Step NUMER in DTM-Vic imposes certain criteria of words and responses to be kept in the final corpus such as minimum word frequency and minimum length of response. Then, step CORTEX helps to merge synonyms and to remove 'stop-words' (i.e., words that are 'useless' to some extent such as 'and,' 'also').

4. Findings

4.1. Descriptive statistics

This section has two main objectives. First, it aims to provide a global picture of the HBs and of the informal HBs in particular in our sample, through describing their characteristics as well as the characteristics of their heads. Second, it presents first descriptive findings from textual analyses.

HBs and informal HBs: Who are they?

Table 1 summarizes the main characteristics of HBs and HB heads in our sample, with a comparison between formal and informal production units. Overall, informal HBs accounted for over two third of the sample and shared a number of common features with the formal HBs, yet the former tended to have more constraints and inferior conditions than the latter. While formal HBs concentrated in trade activities (51 percent), a majority of informal ones focused on services (52 percent).

Table 1. Main characteristics of HBs and informal sector in HCMC (in percentage)

Characteristics	All HBs	Formal HBs	Informal HBs
Characteristics of HBs			
Institutional sector			
Formal	31.67		
Informal	68.33		
Branch of activity			
Manufacturing	20.04	18.21	20.89
Trade	34.7	50.93	27.18
Service	45.26	30.86	51.93
Type of premise			
Unstable (precarious)	25.61	0.93	37.05
Home	51.42	53.09	50.64
Professional	22.97	45.99	12.3
Accounting method			
Professional accounting	1.37	3.72	0.29
Simplified accounting	3.43	8.98	0.86
Personal notes/records	38.63	56.35	30.42
No account	56.57	30.96	68.44
Having wage earners			
No	80.45	61.11	89.41
Yes	19.55	38.89	10.59
Main customer			
Public or parapublic sector	0.49	1.54	0
Enterprise	3.13	4.01	2.72
HB	14.66	15.74	14.16
Individual	81.72	78.7	83.12
Main supplier			
Public or parapublic	6.48	8.91	5.38

Enterprise	16.26	28.71	10.61
HB	73.05	61.06	78.48
Individual	4.22	1.32	5.53
Characteristics of HB heads			
Gender			
Male	43.99	47.22	42.49
Female	56.01	52.78	57.51
Education level			
Primary	24.12	12.42	29.51
Lower secondary	38.82	32.09	40.54
Upper secondary	31.08	41.3	26.36
Tertiary	5.98	11.18	3.58
Apprenticeship			
Technical school	5.92	7.21	5.32
Enterprise or HB	17.46	18.81	16.83
Alone, by practice	67.55	62.07	70.07
Other	9.07	11.91	7.77
Immigration/residence status			
KT1-2 (registered in the same province)	95.21	97.84	93.98
KT3 (temporary registration for 6 months or more	1.86	1.23	2.15
KT4 (temporary registration for less than 6 months	2.94	0.93	3.87
Motivation of creating HB			
Didn't find salaried work (enterprise)	14.82	5.56	19.14
Didn't find salaried work (HB)	8.64	0.62	12.37
To get better income	17.57	24.38	14.39
To be independent	41.51	49.38	37.84
By family tradition	11.38	15.43	9.5
Other	6.08	4.63	6.76
Main reason to choose this activity			
Family tradition	14.96	20.68	12.3
Profession I know	61.68	50.93	66.67
Better profit	6.94	11.11	5.01
More stable returns	9.68	12.96	8.15
Other	6.74	4.32	7.87
Willingness to register your HB			
Yes	21.82	30.25	17.91
No	43.15	1.85	62.32
Do not know	13.7	1.23	19.48
Already registered	21.33	66.67	0.29
Why not register your HB			
Not obligatory			75.5
Do not know			14.04
Other reasons (e.g., complicated, expensive)			10.46
Future perspective			
Yes	41.14	67.08	29.11
No, and think of changing activity	16.93	7.45	21.33
No, but do not think of changing activity	41.93	25.47	49.57

Source: HB&IS HCMC (2009); author' calculations.

Most of the informal sector had no professional premise (88 percent), compared to up to 54 percent for the formal HBs. This is possibly related to the fact that a professional shop in the street could easily draw attention from the authority and thus being more likely to register. Indeed, a professional premise could provide HBs with regular customers and a social network and could be considered as a sign of success, particularly for trade and service businesses.[24] Some illustrative responses could be taken from the in-depth qualitative survey conducted in 2009 and 2010 on 60 HBs in Hanoi and HCMC. Using their own home as premise, which is fairly stable, was the most popular for all HBs; but a large share of informal HBs still operated in unstable premises (37 percent) whilst almost no formal HBs fell into this situation (one percent).

Illustrative cases (Qualitative survey)

Mrs. Trinh (chicken and duck seller):

"Before, I had to pedal around in the city to seek customers so it was more difficult. Nowadays, I can always send to usual customers, since I have a fixed location here."

Mrs. Mai Phi (a tailor):

"I would like to get a bigger premise but to rent premise, I would need more capital. I work at home and it is difficult for my customer to park their motorbike."

Ms. Huong (a hairdresser):

"I have not registered my business as it is currently on a small lane and without signs outside; thus no one came asking me to register. Similar shops on big roads and with shop signs may see local tax officers arrive and ask for registration as well as the payment of taxes."

Source: Qualitative survey, Hanoi and Ho Chi Minh City (2009 and 2010), GSO-ISS/IRD-DIAL (Cling et al. 2010; Razafindrakoto and Nguyen Huu Chi, 2010).

A large majority of the informal HBs maintained no account for their business (68 percent), and the remaining simply used personal notes or records to keep track of their activities. Professional and simplified accounting methods were virtually absent in the informal sector (one percent). The situation was less severe for the formal HBs, with 'personal notes' being the most frequently used form of accounting (56 percent), followed by 'no account' (31 percent). The informal sector seldom employed wage workers (11 percent), while a great proportion of the formal HBs had wage employees (39 percent).

Regarding main types of customers and suppliers, most of both formal and informal HBs sold their goods or services to individuals (79 and 83 percent, respectively) and bought their inputs from other HBs (61 and 78 percent, respectively). Some illustrative cases could be drawn from the qualitative survey in 2009 and 2010. Notably, no informal PUs had a public or para-public institution as their main customer. Actually, the modalities of 'export' or 'import' were added into the questionnaire but no HBs declared either a backward or a forward linkage with foreign markets.

[24] Source: Qualitative survey, Hanoi and Ho Chi Minh City (2009 and 2010), GSO-ISS/IRD-DIAL (Cling et al. 2010; Razafindrakoto and Nguyen Huu Chi, 2010).

These findings suggest that the HBs in general and the informal sector in particular were little integrated into the rest of the economy.

Illustrative cases (Qualitative survey)

Ms. Thu (head of a small business: retail trade):

She was told by a relative that there are premises available for rent and that she should try selling something. She decided to sell a whole array of different things (junk foods, drinks…) and to work also as a hair washer. She bought the commodities either at the marketplace or from vendors who passed by every day. Her customers are people living around, especially young students.

Source: Qualitative survey, Hanoi and Ho Chi Minh City (2009 and 2010), GSO-ISS/IRD-DIAL (Cling et al. 2010; Razafindrakoto and Nguyen Huu Chi, 2010).

Turning to demographic characteristics of HB heads, females were more represented in running HBs, especially in the informal sector. In general, HB heads seemed to have a low level of education and training. Only six percent held a tertiary degree, actually most of them were at secondary educational level (70 percent). Similarly, merely six percent were trained in technical schools, while over two third of HB heads self-trained in practice. Surprisingly, almost all of HBs in our sample were not migrants from other provinces to HCMC. The urban informal sector is sometimes believed as a residual component of the economy that serves as a job provider for migrants from rural areas who could not obtain a job in the formal sector immediately. However, up to 94 percent of informal HBs in our sample actually permanently registered their residence in HCMC.

The main reason of HB heads to create their own business was 'to be independent,' which accounted for 49 percent of formal HBs and 38 percent for informal HBs. A HB head in the HB&IS 2009 survey, for example, said that *"I work as a freelancer, relaxing, upon orders."* Similar responses are: *"(My job is) freelance, I work when I like and I rest when I want"* or *"free business, do not need to depend on anybody."* This result is in line with the main advantage suggested by the qualitative survey in 2009, which is 'autonomy', being their own boss (Razafindrakoto & Nguyen, 2010). For the formal HB heads, the second reason motivating them to set up the business was to generate better income (24 percent), followed by 'family tradition.' In fact, a HB might be an important or even unique source of income for the whole family. By contrast, the informal HB heads opened their own business largely because they were unable to find a salaried work elsewhere (32 percent).

Illustrative cases (Qualitative survey)

Mrs. Hong (seamstress):

"*I am the one who manage the money. I manage my time: there is no pressure, if I want to work, I work, and even if I have a lot of work, I can decide to rest. If I worked for another person, I would be obliged to work to satisfy every customer's demand and I would be constrained to scarify relationships with relatives.*"

Mrs. Thu (head of a small business: retail trade):

"*The greatest benefit of this kind of business is freedom (you could do whatever and whenever you like) and also a bit of pride of being the owner of a shop (not the employee of another person). Yet, what is troublesome is the riskiness of the business. The income is usually unpredictable and does not always permit to cover costs.*"

Mr. Xuan (head of a HB specialized in manufacture of hoarding (advertising boards))

Past job: employee in a state-owned company

"*In fact, I do not earn much from this current business, though my earnings have improved a little compared with what I got from my previous job. Even when the income from the workshop is not very high, it is still an important source of income, though not the major one, for my family. My wife is a lecturer at the university, but we have two daughters who are still going to college.*"

Mrs. Lua (sells rice in the room at the front of the house rented by the family):

"*My husband works as a salesman in a factory (formal sector). My father sells drinks and tobaccos at an improvised post on the road. The current business helps improve my income greatly. Along with my parents' income from their own business, I do contribute much to the whole income of the family."*

Mrs. Huong (hairdresser):

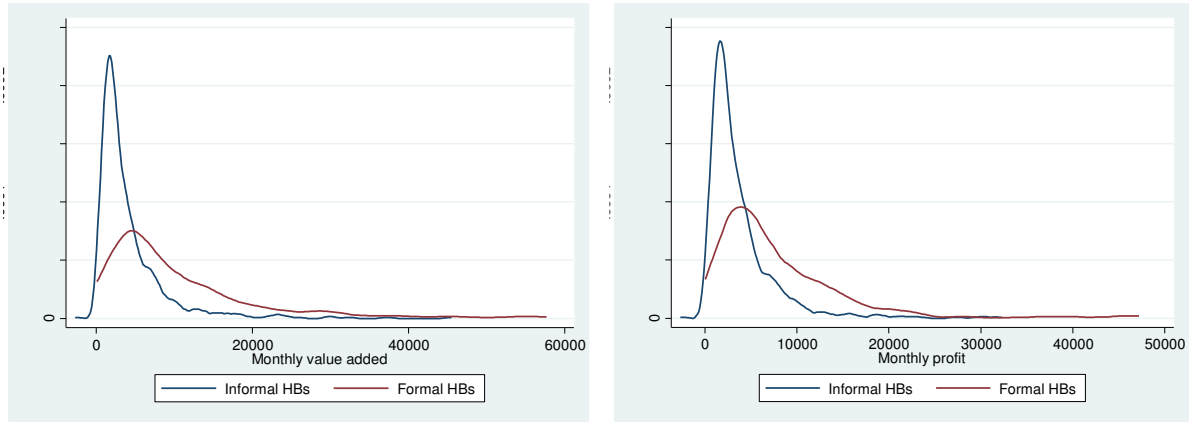
"*In 1984, after completing high school, I decided to look for a work but there was no job opportunity. Then I follow a training to become a hairdresser. In the meantime, I ran a small business (a retail trade, selling different kinds of products). But earnings were very low and unstable. In 1990, I decided to open a hair salon (re-arranging my house). Now, income is higher and more secure than before.*"

Source: Qualitative survey, Hanoi and Ho Chi Minh City (2009 and 2010), GSO-ISS/IRD-DIAL (Cling et al. 2010; Razafindrakoto and Nguyen Huu Chi, 2010).

As improving income was mentioned as a main motivation of workers to run their own HB, it is relevant to take a look at HBs' economic performance. **Figure 1** presents the distribution of HBs according to their monthly value added and profit measured in thousand Vietnam dongs. The performance of HBs in general and in the informal sector in particular was heterogeneous: Most of HBs were positioned near to the subsistence level; however, some HBs managed to realize excellent economic outcomes. Overall, the informal HBs were more present in the lower tail of the distribution than the formal counterparts.

It appeared that the interviewees chose their current activity mostly because of skill constraint (i.e., 'profession that I know'), which constituted 62 percent of HBs as a whole. For example, a HB surveyed in HB&IS 2009 responded to the open ended question that he/she *"has a low level of education, so only makes a small trade to earn a living."* The simplicity of some activities seemed to be a key factor. For instance, a HB head said: *"I do not know what to do, so I run my own motorbike at hand to earn additional income."* Likewise, another interviewee explained that: *"We only need a sewing machine and to know sewing technique in order to do (this job), no worries about the weather."* 'Family tradition' also played a considerable role (15 percent), but it is also related to the professional knowledge, experience and network.

Figure 1. Distribution of HBs According to Economic Performance (in Thousand Vietnam Dongs)



Source: HB&IS HCMC (2009); author' calculations. Note: Some observations with highest recorded values (outliers) are excluded in this histogram for better visual illustration.

Most of the HBs in the informal sector were unwilling to register their business (62 percent) or uncertain of doing so (19 percent). The qualitative survey reveals that some HBs preferred to stay informal to avoid paying taxes, in line with the 'legalist' or 'orthodox' approach proposed in de Soto (1989).

Some illustrative cases (Qualitative survey)

Nguyen Thi Phuc (hairdresser):

"*I do not see any advantages from business registration. Also, I have never met any difficulties with my current (informal) status.*"

Pham Thi Hang (seller of ice blocks and cubes):

"*I do not know (any advantages or disadvantages) as I have not registered. If allowed to choose voluntarily, I would not choose to register my business.*"

Truong Minh Quang (producer of steel doors):

"*There are no benefits from registering. The disadvantage is to have to pay taxes every month and to be disturbed by local offices asking for donations.*"

Do Thi Thanh (lawyer):

"*Formality brings no benefits. The disadvantage is to have to use invoices in your business, to have to pay taxes. If my business is not registered, I have the freedom to decide to work or to have a break. If it is registered, then I must suffer the pressure about the rental, the salaries for employees and the taxes.*"

Source: Qualitative survey, Hanoi and Ho Chi Minh City (2009 and 2010), GSO-ISS/IRD-DIAL (Cling et al. 2010; Razafindrakoto and Nguyen Huu Chi, 2010).

Notably, three quarters of unregistered HBs believed that registration was not obligatory for their business, and many of them just did not know (14 percent). There seemed to be a vague

awareness about the regulations regarding registration among HBs. The responses to open-ended questions in HB&IS 2009 show that many HBs thought they did not have to register because their business was so small. This was also the case in the qualitative survey, together with other justifications such as manual work, family tradition, old age, etc. In fact, according to the law, local authorities shall set levels of low income in their respective localities, under which the HBs would not have to register their business. Yet, these levels of low income are not publicly announced. What is more, since the real income of HBs was difficult to check, the decision regarding which HBs are exempted from registration easily becomes inaccurate.

Some illustrative cases (Qualitative survey)

Do Thi My Trinh (bus rental):

"*I think that to register, you need to visit the Department of Planning and Investment, yet I am not sure about the whole procedure. My own business is non-registered because of minuscule scale (just having one bus).*"

Do Minh Hoang (head of a civil house-building unit):

"*I do not register because I see that it is unnecessary; my current job is spontaneous.*"

Nguyen Thi Kim Hong (tailor):

"*I think that my business involves no financial capital, only manual work. Only if I expand my business, I would be obliged to register."*

Ngo Quoc Huyen (seller and transporter of construction materials):

"But it is just because I conduct my business only around this neighborhood, because the distance over which we transport the materials is rather short. Was it longer, I would have registered."

Nguyen Cong Thinh (producer of copper products):

"My business is conducted on a basis of family tradition and at home; it is not similar to 'having a shop' thus no obligation to register and pay tax. As a matter of fact, authorized officer of this neighborhood is well aware of my family's business without asking me to register it."

Nguyen Thi Thanh (motorbike cleaner):

"Someone arrived to check my business but they say that I am over 60 now so I do not have to register or to pay taxes. Even the shop over there, they earn several hundreds of VND a day without registering."

Source: Qualitative survey, Hanoi and Ho Chi Minh City (2009 and 2010), GSO-ISS/IRD-DIAL (Cling et al. 2010; Razafindrakoto and Nguyen Huu Chi, 2010).

Merely ten percent of unregistered HBs put forward reasonable explanations such as 'too complicated steps,' 'too expensive,' 'do not want to collaborate with the State' or 'registration in progress.' Actually, the qualitative survey uncovers various experience of HB heads in registering their business. Some underwent rather simple procedures, while others complained about the complexity and delay of the process.

Some illustrative cases (Qualitative survey)

Ngo Xuan Quynh (seller of gas and gas cooker):

He went through a multi-step procedure. Before formally registering his business, he had to undergo some procedures related to the business line of selling liquidized gasoline including: certificate of compliance with fire-fighting regulations issued by Fire-fighting Brigade Office; certificate of eligibility for selling liquidized gasoline issued by the Department of Trade; certificate of compliance with safety and security requirements issued by the Office of Police responsible for social security – under the Police of Hanoi city. Those certificates were again the required documents for Mr. Quynh to fill in registration dossier. Due to change in the geographic division within his residing neighborhood, he had to re-register his business in 2004 after the first registration in 2001. Regarding time and efforts taken to complete the registration procedures, he says:

"I do not remember exactly, but I had to visit the local authorities approximately 5 or 6 times and I had to wait for about half a month."

When he registered his business again in 2004, Mr. Quynh met some difficulties due to unclear and inconsistent guidance from the in-charge officer. He expressed a bit anger:

"When I applied for a new registration certificate, I originally thought it was very simple. I just thought that I needed to submit an application, no need for special acquaintances. But the officer of the People Committee of the district asked me to make two sets of dossier. When I had prepared the two sets, then he said that three sets are needed. Yet, as I had got one more copy of the dossier and presented all of them to another officer, this officer then said that two were enough. That is what I could not understand; the bureaucratic procedures are difficult to understand."

Source: Qualitative survey, Hanoi and Ho Chi Minh City (2009 and 2010), GSO-ISS/IRD-DIAL (Cling et al. 2010; Razafindrakoto and Nguyen Huu Chi, 2010).

Regarding perception about future perspective of HBs, there was a clear distinction between formal and informal production units. While over two third of the former kept optimistic attitude toward their prospect, most of the latter thought that there would be no future for a business like theirs (71 percent); nevertheless, half of them did not think of changing activity. This question indirectly reveals the generally low job satisfaction and limited dynamics of HBs in the informal sector, possibly due to their skill and capital constraints, or 'self-sufficient' psychology, or preference to stay independent.

Some illustrative cases (Qualitative survey)

Ms. Hang (ice seller):

Ms. Hang was born in Hanoi and has been living for more than 40 years with her family. Though her parents were both workers at nearby factories, she did not find a job as a worker because of lack of education and qualifications. When she graduated from high school more than 20 years ago, she started selling ice.

"*The reasons why I decided to undertake this activity are: first, need for capital is small, not substantial; second, no particular skills or tools are required and I have been used to it for years; thirdly, I could manage this business by myself and compared to other small business activities (e.g., selling food in marketplaces), incomes are acceptable albeit low. For more than 20 years, I did think of changing activity to sell other products, but then due to lack of capital and premises, I refrained from trying*".

Ms. Tham (shrimp cake seller, for 15 years)

"*Before, I did not manage to find work and this activity helped me to improve my living conditions. But I would like to change because the income is low and the working conditions are so hard and there is no prospect to develop this type of activity. If I manage to find another type of work, with more secure income, I will change immediately.*"

Source: Qualitative survey, Hanoi and Ho Chi Minh City (2009 and 2010), GSO-ISS/IRD-DIAL (Cling et al. 2010; Razafindrakoto and Nguyen Huu Chi, 2010).

First insights from the textual data of HB&IS 2009

The textual data mobilized for the analysis are the precise answers provided by the head of HBs to the two open-ended questions.[25] They reveal copious information that not only directly answers the questions asked, but also provides explanations, inferences and attitudes of the HBs. One HB, for example, put forward difficulty due to price increase, after explaining its causes and mechanisms: *"Gold price increases, gasoline price goes up day by day so goods prices follow the trend."* Another HB considered that goods in their grocery store were impossible to sell because their main customers were manual workers, who were unemployed at the moment. Meanwhile, some HBs reported very job-specific problems such as: *"There was once a customer who ate our sweetened porridge and had a bellyache, she laid the blame on us and scolded noisily."* They even suggested policies to the government: *"Cheap Chinese goods make customers tend to replace rather than to repair, there should be regulations to reduce traffic fees, stabilize petrol prices in order to enhance the transportation industry."* These kinds of information are hardly available in a quantitative survey.

Table 2. Numerical Coding of Texts: A Summary

	Disadvantages	Advantages
Total number of responses	505	433
Total number of words	5188	3498
Number of distinct words*	867	585
Percentage of distinct words	16.7	16.7
Frequency threshold	2	2
Kept words	4551	3045
Distinct kept words	336	229

*Source: HB&IS HCMC (2009); author' calculations. Note: * A single word might be used many times, thus multi-counted.*

[25] The two open-ended questions were analyzed in parallel but separately to avoid misinterpretation (given the fact that the analysis presented is a first exploratory phase), but they could be combined in further steps of the analysis.

Basic statistics resulting from numerical coding of texts are provided in **Table 2**. We report the number of responses (i.e. observations or heads of HBs who have provided answers) to each question and count the total number of words and distinct words contained in the responses. We then remove all the words with frequency equal or less than two, as these words contribute little to the text corpus. **Table 2** shows that the response rate of the 'disadvantage' question is significantly higher than that of the 'advantage' question (72.2% versus 61.9%). Non-responses appear to be random, as most of the key variables (i.e., HB size, value-added, educational attainment of HB heads) are quite similar between the respondents and the non-respondents (**Table 19** in Appendix). **Table 2** also indicates that the number of words and distinct words used by HBs to present their disadvantages is about 50 percent higher than for advantages. These figures, to some extent, reflect the low satisfaction of working in the informal sector. This result is consistent with the findings of Razafindrakoto et al. (2012) based on a scale question about their job satisfaction level.

The HB&IS survey 2009 also includes a close-ended question asking for the problems or difficulties faced by HBs and their degree of severity. One advantage of having both types of question (i.e., close-ended and open-ended) for a similar purpose is the ability to extract as much information as possible from the interviewees. **Table 3** shows that some HB heads said "no" when being asked in the close-ended question, but did reveal something when they were allowed to express in their own words, and vice versa. In general, difficulties prevailed (more than half of HBs had at least a problem), but most of them were minor or moderate (only 16 percent of informal HBs had at least a major or severe problem).

Table 3. Number of responses to the close-ended question versus to the open-ended question concerning the problems and disadvantages of informal HBs

	Responded to 'disadvantage' question		
Close-ended question	No	Yes	Total
Having at least a problem			
No	120	192	312
Yes	74	313	387
Having at least a major/severe problem			
No	172	416	588
Yes	22	89	111
Total	194	505	699

Source: HB&IS HCMC (2009); author' calculations.

Turning back to the open-ended question, a closer look at the responses gives us a list of the most frequent words used by HBs (**Table 4**). In terms of advantages, 'home' mattered the most. This is likely to imply a premise at home, since 'premise' and 'location' also appear very frequently. Two popular explanations can be put forward: either a HB owner had a premise at home and she saved 'rent,' or she was happy working at home to have 'time' taking care of her 'family.' 'Customers' was the second most important advantage cited by informal HBs, and they could be the 'regular' customers. 'Stability' is also one of the main advantages revealed, which reflects the risk aversion of these HBs. As mentioned previously, a HB might be the main or a

considerable source of 'income' for the whole family. Another advantage of running an informal HB is low requirement of initial 'capital,' low 'cost' and no (or little) 'tax' burden.

Table 4. Most Frequent Words Used by Informal HBs

Advantages	Frequency	Disadvantages	Frequency
Home	78	Customers	102
Customers	67	Prices	79
Stable	57	Capital	56
Premise	54	Income	47
Location	50	Location	35
Rent, hire	46	Road, street	34
Time	41	Competition	32
Family	37	(Transport) Vehicle	32
Income	34	Money, cash	30
Familiar, regular	29	Borrow, lend	28
Money, cash	26	Home	26
Profession	26	Profit	25
Capital	25	(Un)stable	25
Cost	21	Small	24
Near, close	20	Lane	23
Tax	17	Rain	22

Source: HB&IS HCMC (2009); author' calculations.

It might be interesting to compare the main difficulties declared by HBs in the close-ended question with those in the open-ended question as listed above. First, this comparison could reveal some important problems that may be absent or underrepresented in one of the two questions. Second, this could give an insight into the methodological problem regarding question choice (interesting examples are provided in Piau, 2004).

As suggested by the close-ended question (**Table 5**), 'competition' is the most predominant difficulty encountered by the HBs. 'Lack of customers,' 'access to loan,' 'premise, space,' and 'cash flow' come in the following places. Reasonably, these five issues also appear in the most frequent words in textual analysis.[26] However, the order is different. 'Customers' instead of competition, becomes their primary concern. Interestingly, 'prices' emerges as a major issue in the open-ended question, whereas it is faint in the close-ended question. This may be due to the fact that 'prices' is not explicitly mentioned in the latter. It is possibly grafted into 'inflation and exchange rate' category, which might appears as an abstract concept to informal HBs. Moreover, this category is positioned toward the end of the list, suffering from the 'primacy effect,' i.e., when the question is long, respondents tend to choose the first categories (Schuman & Presser, 1996). Further, the textual data reveal some important difficulties facing informal HBs that are not captured in the quantitative part of the survey such as 'rain,' (i.e., weather, especially for itinerant street vendors), 'road, street' (e.g., the urban police expelling businesses settled on the pavements).

[26] 'Premise, space' is related to words such as 'location,' 'road, street,' 'home,' 'hamlet' cited in texts.

Table 5. The Main Difficulties Faced by Informal HBs

The most difficulty faced	Frequency	Percentage
Competition	119	30.75
Lack of customers	57	14.73
Access to loan	55	14.21
Premise, space	49	12.66
Cash flow	25	6.46
Machine, equipment	21	5.43
Material supply	18	4.65
Crime, theft, disorder	13	3.36
Inflation, exchange rate	11	2.84
Transportation	9	2.33
Other	7	1.81
Access to land	3	0.78

Source: HB&IS HCMC (2009); author' calculations.

One might notice that some words could indicate both advantage and disadvantage (the colored words). This is primarily because a noun just neutrally presents an issue, a business aspect, and whether this aspect is in good condition or not would determine if it can be considered as advantageous or disadvantageous. The following box quotes some cases illustrating how 'home,' 'location' and 'customers' could be advantages for some HBs and disadvantages for others.

'Home,' 'customer' and 'location' could be either advantages or disadvantages: Some illustrative cases

Advantages

"Business location is fairly advantageous, premise at home so do not need to rent."

"Sell at home so convenient for family activities."

"Sell at home, can take care of the house and of the children at the same time."

"Can do housework and business simultaneously, I sell when there are customers, I do housework when there are no customers."

"Close from home, convenient for travel and business."

"Hire a good location to do business, customer quantity is stable."

"Good revenue, favorable location for traffic vehicles, for customers to come to play."

"Sell at home, have regular/familiar customers."

"Have a propitious location to do business, I have sold for many years thus I am experienced and have regular/familiar customers."

Disadvantages

"Previously I sold in a marketplace very well, now I am selling at home, because the street is deserted, earnings are not much."

"Premise is small and narrow, teaching at home makes communication with customers limited."

"Surrounding environment is polluted due to the bottle ('ve chai') sound near my home."

"Doing business far away from home, no one cares for my small children."

"Selling far away from home, petrol-consuming."

"Too far from home, hard to monitor and manage."

"In rainy seasons, I usually have few customers because of selling on the road or pavement."

"In a lane so there are few customers, customers are only acquaintances."

"Location is in a lane so the number of customers is limited, need more capital."

"Customer quantity is few because of a bus line, only able to carry the customers nearby."

"Have to hire a premise, electronic appliances are cheap so if they break down, the users just throw them away and buy new ones, therefore the repair service does not go well."

"Competition is too harsh, capital is little, hard to keep customers."

"Delivering water right to customers' houses so whenever customers call, I have to carry water to them."

Source: HB&IS HCMC (2009). Author's translation and compilation.

4.2. Correspondence Analysis on the open-ended questions

There are two distinguished groups of variables involved in this analysis. *Active variables,* which are in this case the first two open-ended questions, will contribute to the performance of correspondence analysis and explain the *factors* derived from the analysis. Since it would be confusing for the interpretation if we mix advantages with disadvantages, correspondence analysis is performed separately for each question.[27] Besides, *supplementary variables*, which are in this case categorical variables constructed from the quantitative data, are added into the analysis to provide more informative results.

The explanatory power of the principal components is reflected through their corresponding eigenvalues as well as percentage in the total variance (**Table 6**). The more the

[27] The correspondence analysis for the combined corpus of advantages and disadvantages is briefly presented at the end of this section 4.2 for reference.

dimensionality of the original data, the lower eigenvalues and percentages.[28] Acknowledging the number of distinct words contained in 'advantages' and 'disadvantages,' which is equivalent to the number of textual dimensions (229 and 336, respectively), the figures below are relatively large. The first factor of 'advantages' for instance, accounts for nearly 2% of the total variance, whose explanatory power equals to that of four and a half variables in the original data set.

Table 6. The First Five Eigenvalues of Correspondence Analysis

	Advantages		Disadvantages	
Factor	Eigenvalue	Percentage	Eigenvalue	Percentage
1	0.5848	1.96	0.5114	1.39
2	0.5597	1.88	0.4925	1.34
3	0.5311	1.78	0.4878	1.32
4	0.5093	1.71	0.4578	1.24
5	0.494	1.66	0.4537	1.23

Source: HB&IS HCMC (2009); author' calculations.

Hereafter, we consider only the first two most variance-explained factors. The interpretation of principal axes is based on the words that have extreme contributions to their creation (**Table 7**, **Table 8** and **Table 9**).[29]

Advantages

The meaning of factorial axes formed by HBs' advantages is illustrated in **Table 7** and **Figure 2**. In **Table 7**, the words with highest absolute values of coordinate are those contribute the most to the interpretation of the axes.

The first factorial axis clearly differentiates *social* and *non-social* logics. On the negative side of the axis, informal HBs take advantage of their social networks such as local authority, friends and family. On the positive side, informal HBs find the advantages right from the nature of their business (i.e., 'small' scale, no or little 'tax' duty, 'premise,' 'initiative,' 'time' and so on).

The second factorial axis contrasts *maximizing* against *satisfying* logics. Some informal HBs, on the one hand, actively 'make use of' as many opportunities as they can. This is demonstrated through words such as 'try,' 'initiative' and economic advantages that they mention like 'tax,' 'premise,' 'production,' 'expenses.' On the other hand, other informal HBs seem to lower their ambition to achieve job satisfaction. These people choose less demanding jobs which are 'easy,' 'light,' 'simple' and may be appropriate to their 'health' and 'age,' in order to earn a just 'sufficient' and 'acceptable' living.

[28] The Kaiser-Gutman rule for correspondence analysis indicates that we should keep the factorial axes whose eigenvalues greater than average eigenvalue. This is equivalent to eigenvalue-greater-than-one criterion for PCA (Guttman, 1954; Kaiser, 1960, 1970)

[29] We eliminate several words that can be considered as 'outliers,' i.e., not belong to the systematic meaning of the axes.

Table 7. Selective Points with Extreme Contributions to the First Two Principal Axes: Correspondence Analysis on "Disadvantages" of Informal HBs

Axis 1			Axis 2		
Words	**Translation**	**Coordinate**	**Words**	**Translation**	**Coordinate**
Giúp đỡ	help	-11.85	Giúp đỡ	help	-15.66
Chính quyền	authority	-10.23	Chính quyền	authority	-14.85
Địa phương	locality	-8.46	Thuế	tax	-10.50
Bạn bè	friends	-6.51	Cố gắng	try	-4.11
Gia đình	family	-5.72	Kinh doanh	business	-2.34
Quen biết	acquainted	-5.26	Mặt bằng	premise	-1.58
Cố gắng	try	-4.47	Tận dụng	make use of	-1.51
Thu nhập	income	-4.27	Sản xuất	production	-1.33
Quen	familar	-3.19	Kinh phí	fund, expense	-1.26
Nguồn	sources	-3.01	Chủ động	initiative	-1.21
Thời gian	time	1.80	Tuổi	age	1.33
Chi phí	costs	1.85	Tạm	acceptable	1.49
Chủ động	initiative	2.62	Con	children	1.67
Địa điểm	location	2.83	Ổn định	stable	1.79
Nhà	home	3.08	Sức khỏe	health	1.86
Kinh doanh	business	4.43	Sống	live	1.91
Thuê	hire	4.57	Đủ	sufficient	1.99
Mặt bằng	premise	5.15	Thời gian	simple	2.04
Nhỏ	small	6.19	Nhẹ nhàng	light	2.05
Thuế	tax	12.82	Dễ	easy	2.35

Source: HB&IS HCMC (2009); author' calculations.

I then investigate the link between HBs' advantages with the motivation of heads of HB to set up their business and to choose their business activity, by projecting motivation categories on the plain of these two factorial axes.

Hypothesis: HBs that are located in the non-social and maximizing sphere are more likely to be established as a voluntary choice. In contrast, HBs which are positioned towards the social and satisfying sphere are more likely to be created due to a constraint.

As expected, **Figure 2** suggests that HBs which were created voluntarily (i.e., to get better income, to have higher profit than other activities) rather than due to a constraint, either employment constraint (i.e., because the HB head did not find a job in another HB) or expertise constraint (i.e., because this is the profession that they know) are located towards the bottom right of the dimensional space formed by the two first principal axes. In other words, 'voluntary HBs' are less dependent upon personal networks such as family, neighbors and friends in their business operation. They also take advantage of economic opportunities to develop their business.

Disadvantages

Table 8 shows the words that contribute to the interpretation of the first two principal axes generated from disadvantages of informal HBs. These words are also projected on the factorial plane of the first two principal axes (**Figure 3**). On the same plane, **Figure 4** projects the categorical variables to further explain the characteristics of HBs located in different positions.

Table 8. Selective Points with Extreme Contributions on the First Two Principal Axes: Correspondence Analysis on "Disadvantages" of Informal HBs

Axis 1			Axis 2		
Words	**Translation**	**Coordinate**	**Words**	**Translation**	**Coordinate**
Vốn	capital	-7.32	Sớm	early	-16.33
Tài chính	finance	-7.01	Dậy	get up	-15.89
Vay	borrow	-4.97	Thức	awake	-13.14
Hỗ trợ	support	-4.88	Sang	morning	-10.72
Tay nghề	workmanship	-4.48	Vất vả	strenuous	-6.23
Mở rộng	enlarge	-4.43	Vốn	capital	-5.69
Diện tích	area, surface	-4.26	Tài chính	finance	-3.76
ổn định	stable	-3.04	Vay	borrow	-3.06
Quy mô	scale	-3.02	Gia đình	family	-3.04
Máy móc	machines	-3	Sản xuất	production	-2.94
Kinh tế	economy	-2.85	Mở rộng	enlarge	-2.76
Mặt bằng	premise	-2.84	Diện tích	area, surface	-2.53
Giải tỏa	land clearance	-2.66	Già	old	-2.4
Vật giá	prices of materials	-2.65	Chồng	husband	-2.04
Địa điểm	location	-2.52	Con	children	-1.82
Mất	lose, lost	2.45	Thị trường	market	2.07
Trung Quốc	China	2.51	Chất lượng	quality	2.2
Bụi	dust, dusty	2.81	Bụi	dust, dusty	2.25
Xăng	petrol	3.38	Siêu thị	supermarket	2.28
Xe ba bánh	tricycles	3.41	Nắng	sunny	2.57
Kẹt	(traffic) jams	3.74	Nước	water	2.57
Chất lượng	quality	3.85	Mưa	rain	2.68
Cấm	ban	4.36	Khách hàng	customers	2.79
Vất vả	strenuous	4.37	Dọn dẹp	tidy up	2.83
Khách hàng	customers	4.54	Buýt	bus	2.92
Đường	road, street	4.94	Kẹt	(traffic) jam	3.53
Sang	morning	8.31	Xăng	petrol	4.09
Thức	awake	9.4	Xe	vehicle	4.65
Dậy	get up	10.87	Đường	street, road	5.36
Sớm	early	11.46	Giá	prices	6.64

Source: HB&IS HCMC (2009); author' calculations.

The first factorial axis distinguishes *entrepreneurial* from *personal* perspectives. Those who are located towards the negative side of the axis are more likely to raise their concerns in an entrepreneurial perspective. These informal HBs stress more economic problems such as 'capital,'

'finance,' 'scale' and 'prices of materials.' In contrast, those who are positioned towards the positive side of this axis tend to put forward their difficulties in a personal perspective. Their disadvantages, for instances, are the fact that they have to 'get up' 'early' in the 'morning' to work, which is 'strenuous' as far as they are concerned.

The second factorial axis separates out *internal* and *external* disadvantages: (i) those that come from the informal HBs themselves and that they can actively improve the situation (e.g., 'strenuous,' 'finance,' 'family' and 'production'); and (ii) those that are external and almost cannot be changed by HB's own actions (e.g., 'street,' 'petrol,' 'dust,' 'rain' and so on).

Hypothesis: HBs positioned towards *entrepreneurial* and *internal* directions would be the *high-end* informal HBs and vice versa. They probably operate on a larger scale, have better economic performance and more stable working conditions.

As a validation of this hypothesis, we project the significant quantitative modalities onto this plane (**Figure 4**). Impressively, we observe an evolution of informal HBs from the top right to the bottom left that: scale is enlarged (from 'size=1,' 'size=2' to 'size3-6'), business turnover increases (from the first quintile to the fifth quintile), business premise is stabilized (from unstable premise to premise at home), their attitude toward their business becomes more optimistic (from thinking that their HB has no future to the reverse), and so on.

This phenomenon probably stems from the fact that different HBs in different stages of development face their respective obstacles. Assume that all HBs concern about their operating conditions/environment. For an own-account worker ('size=1') who wanders the streets or sits on the pavement ('premi_unstab'), for instance, her environment is the streets, including the weather, dust, vehicles and local authority expelling vendors like her. Meanwhile, the business environment of a big, stable shop should be less affected by natural conditions. Instead, their owner could concern more about the production and market conditions such as 'capital,' 'finance,' 'scale' and 'prices of materials.' Moreover, the disadvantages mentioned could also reflect the business attitude and future perspective of HB owners. Some disadvantages are truly obstacles to a normal operation of HBs such as bad weather and traffic jams for those without a stable premise. Others are considered 'disadvantages' simply because the interviewees have intention to enlarge, to develop their business. More ambitious respondents tend to be those who are currently running a well-performed HB.

A broader view on both informal and formal production units (**Table 9**, **Figure 4** and **Figure 5**) further elucidates the heterogeneity of the urban HBs. It should be noted that the majority of the analyzed sample of non-farm HBs is informal (68%) and that medium, large and formal enterprises are excluded from the HB&IS Survey (Cling et al., 2010b). For this reason, these results obtained here are very relevant to the informal sector.

The first factorial axis is mainly on disadvantages which affect *personally* each HB: From those who face individual constraints (e.g., 'get up' 'early' in the 'morning') to those who are more integrated into the market economy and put forward business and economic constraints (on the supply side) in terms of 'capital,' 'premise,' 'production,' 'interest rate' and 'machines,' for examples.

The second factorial axis appears to be more on disadvantages faced *globally* by HBs: from economic and market-related difficulties (on the demand side) such as 'prices,' 'competition,' 'market,' and 'taxes' to non-economic, non-market factors (e.g., 'weather,' 'rain,' 'season,' 'flooded').

Table 9. Selective Points with Extreme Contributions on the First Two Principal Axes: Correspondence Analysis on "Disadvantages" of HBs

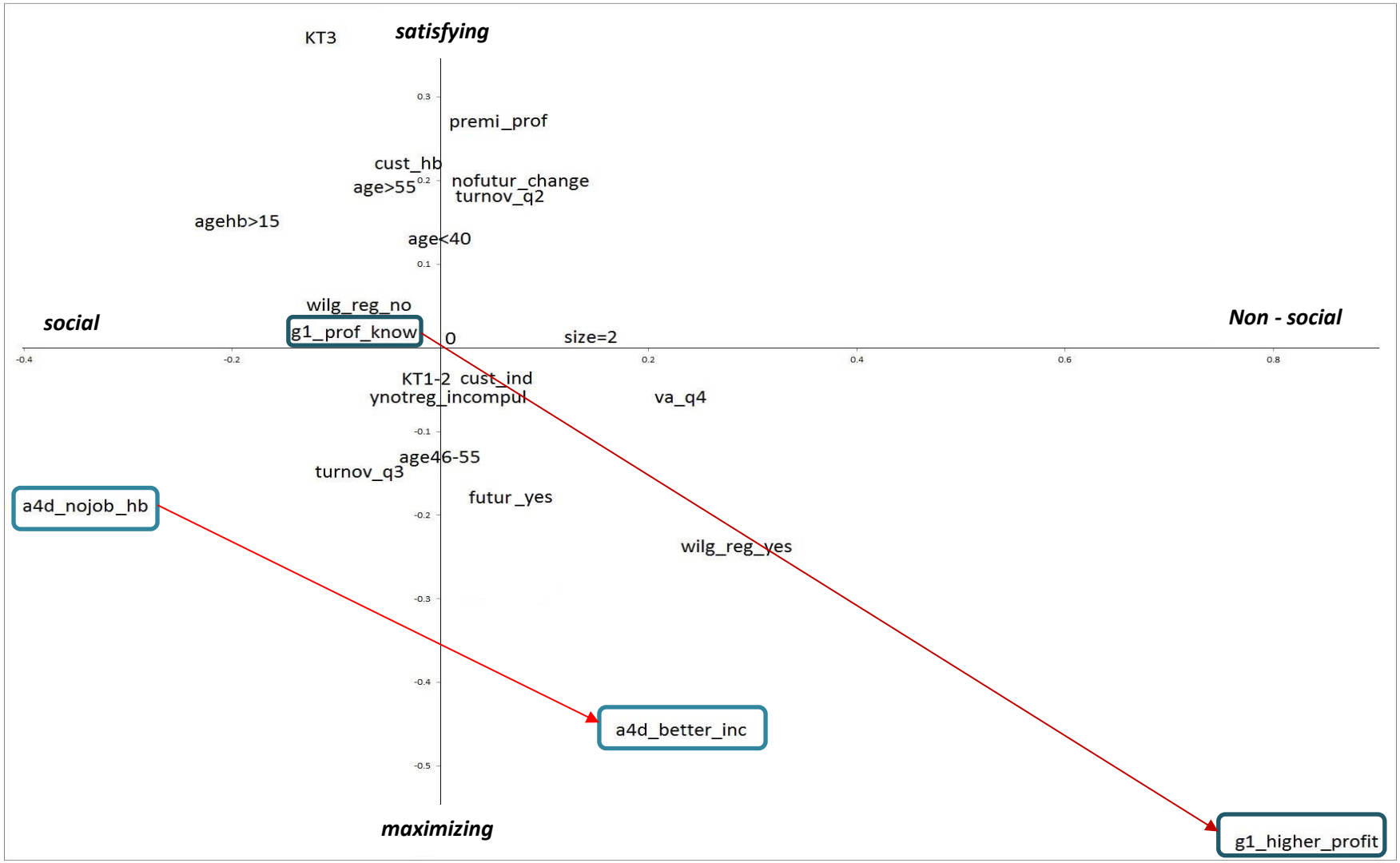
Axis 1			Axis 2		
Words	**Translation**	**Coordinate**	**Words**	**Translation**	**Coordinate**
Sớm	early	-13.1	Giá	prices	-10.47
Dậy	get up	-12.57	Cạnh tranh	competition	-7.22
Thức	awake	-10.44	Gia công	processing	-6.5
Sang	morning	-9.74	Đối thủ	competitors	-4.06
Khách	customers	-5.99	Thị trường	markets	-3.31
Xe	vehicle	-5.69	Lợi nhuận	profit	-3.15
Đường	street, road	-4.97	Điện	electricity	-3.05
Vất vả	strenuous	-4.51	Chi phí	costs	-2.85
Cấm	forbid	-4.06	Tiêu thụ	consume	-2.69
Quản lý	manage	-3.67	Dân cư	inhabitants	-2.67
Bụi	dust	-3.42	Thuế	tax	-2.6
Xe ba bánh	tricycles	-3.23	Tay nghề	workmanship	-2.51
Xăng	petrol	-3.16	Sản phẩm	products	-2.45
Kẹt	(traffic) jam	-3.05	Doanh thu	revenue	-2.39
Lề	edge	-2.52	Chất lượng	quality	-2.29
Chật hẹp	narrow	3.14	Bụi	dust	2.96
Quy mô	scale	3.65	Xe ba bánh	tricycles	3.61
Máy móc	machines	4.09	Nước	water	4.02
Lãi suất	interest rate	4.1	Cấm	forbid	4.37
Địa điểm	location	4.32	Nắng	sunny	4.6
Sản xuất	production	4.44	Dọn dẹp	tidy up	5.26
Ngành nghề	industry	4.5	Vất vả	strenuous	5.38
Diện tích	surface area	4.7	Sáng	morning	5.67
Mặt bằng	premise	5.42	Ngập	flooded	6.42
Di dời	move	5.56	Đường	street, road	6.5
Vay	borrow	6.99	Mùa	season	7.37
Mở rộng	enlarge	7.31	Dậy	get up	8.98
Tài chính	finance	8.4	Sớm	early	9.42
Kinh doanh	business	9.03	Trời	weather	13.35
Vốn	capital	10.81	Mưa	rain	15.21

Source: HB&IS HCMC (2009); author' calculations.

As a further attempt to visualize the heterogeneity of the HBs, we project the significant words and quantitative modalities onto the plane of the first two principal axes above. It is expected that HBs who tend to put forward economic constraints, both on supply and demand sides, are more likely to be *high-end* HBs. Reversely, HBs who complain about non-economic and individual difficulties could generally belong to the *low-end*.

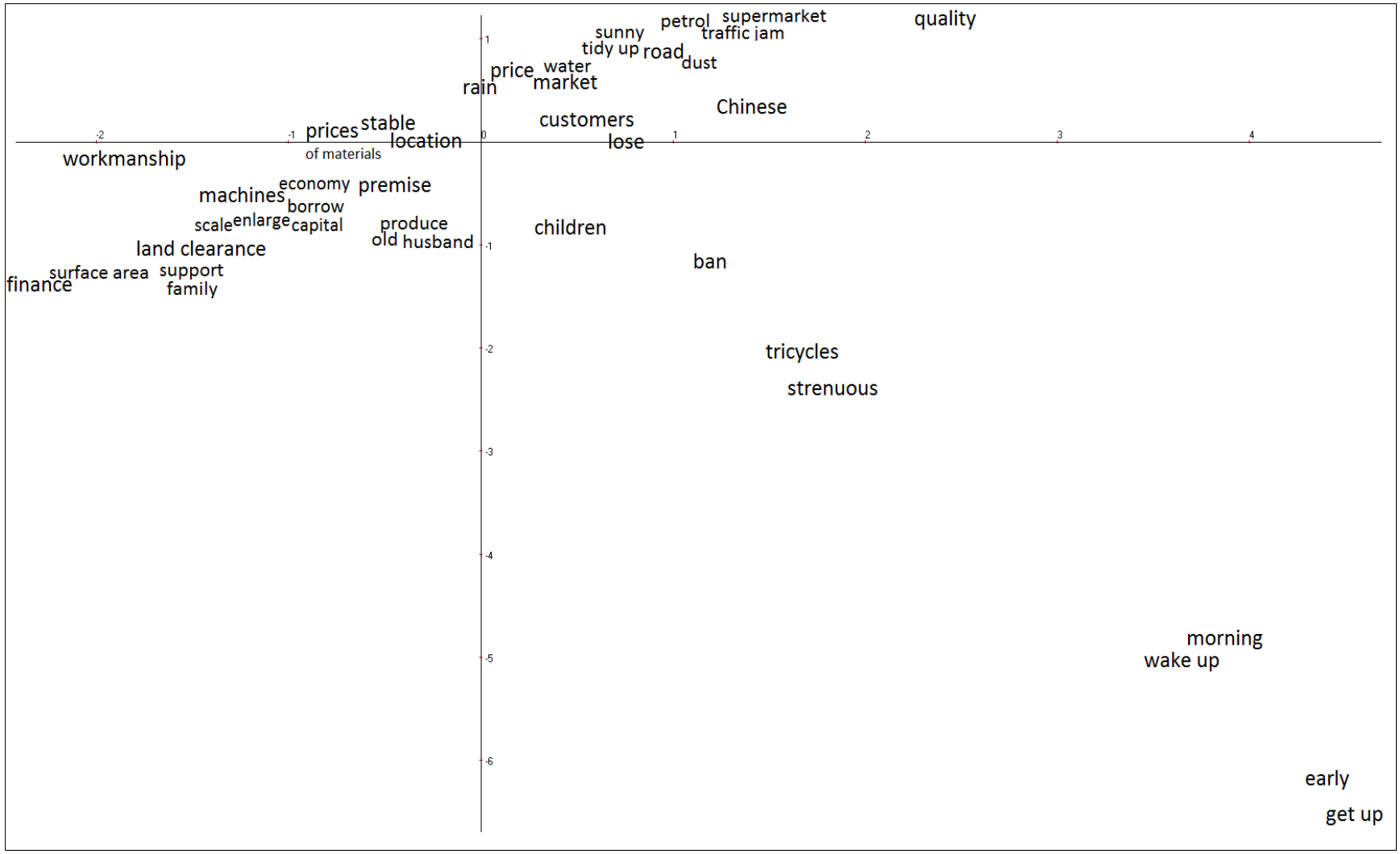
Figure 6 is eye-catching: From the top left of the figure, we have the *low-end,* i.e., the most precarious, the poorest, the least educated, the less integrated into the economy; to the right-bottom of the figure with the *high-end* businesses, i.e., the biggest, the best-performed, the more educated, the more embedded in the economy and so on. This figure is the clearest illustration of our argument that the informal sector is rather a continuum, than a homogeneous sector, or a dual segmented phenomenon. Business size, for instance, gradually grows from 'size=1' in the top left quarter to 'size3-6' and then to 'size>6' in the far bottom right of the plane. School attendance of HB head, in the same manner, escalates from primary to secondary school, and finally high school.

Figure 2. The Space of Informal HB Characteristics (First Factorial Plane of Correspondence Analysis on Advantages)



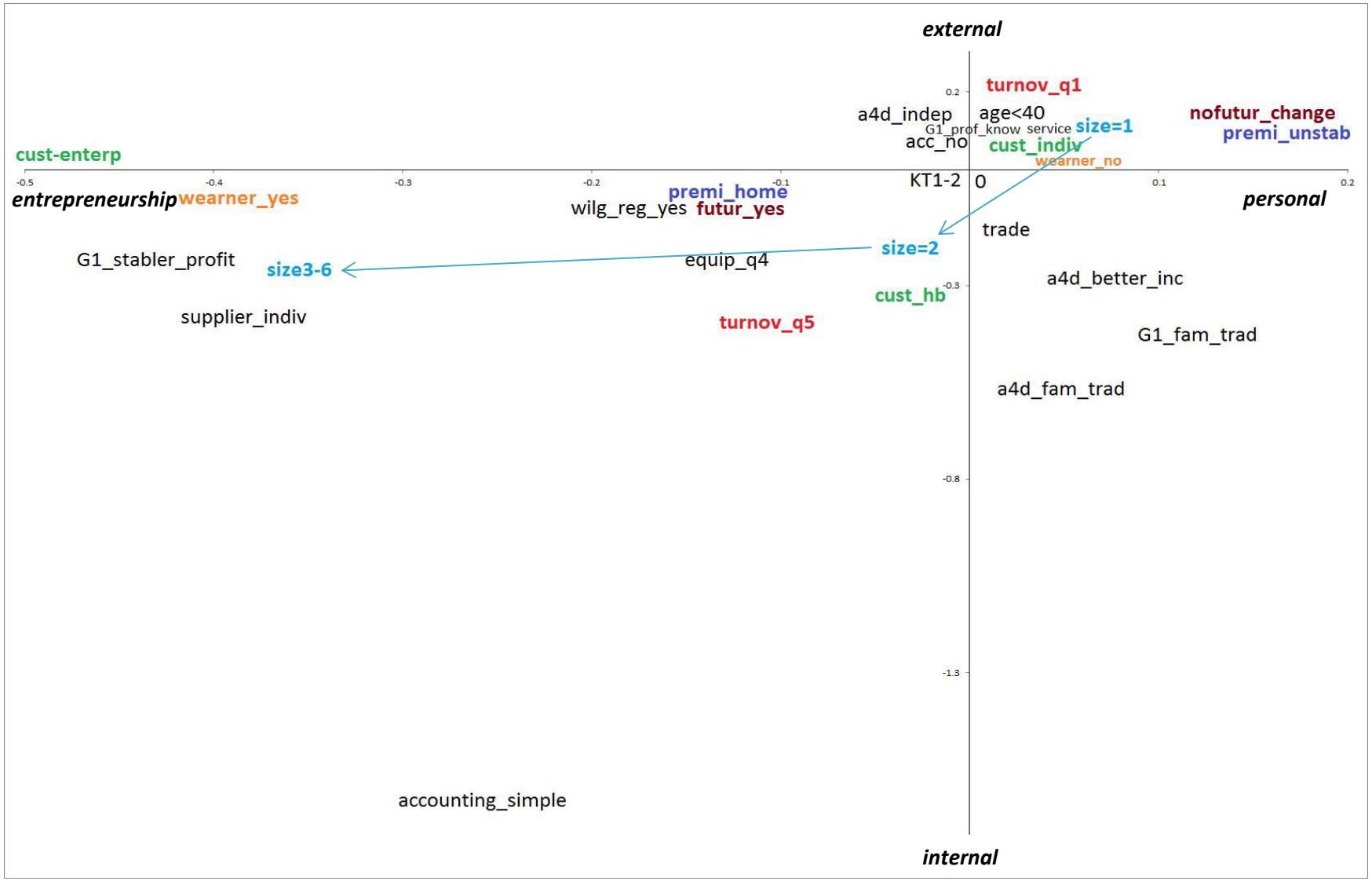
Source: HB&IS HCMC (2009); author' calculations. Note: The explicit meaning of the categorical variables is provided in Table 18 of Appendix.

Figure 3. The Space of Disadvantages Put Forward by Informal HBs (First Factorial Plane of Correspondence Analysis)



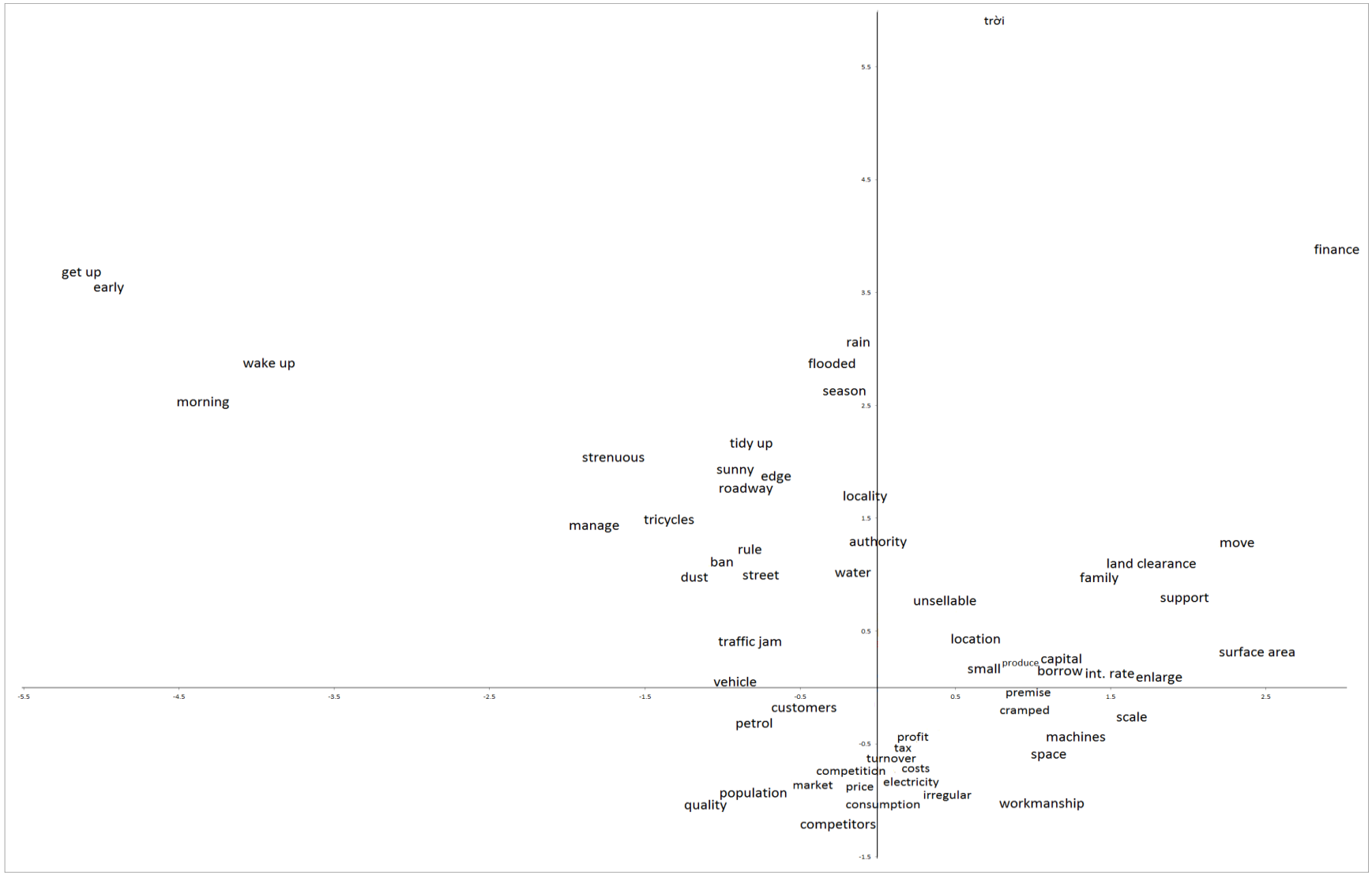
Source: HB&IS HCMC (2009); author' calculations.

Figure 4. The Space of Informal HB Characteristics (First Factorial Plane of Correspondence Analysis on Disadvantages)



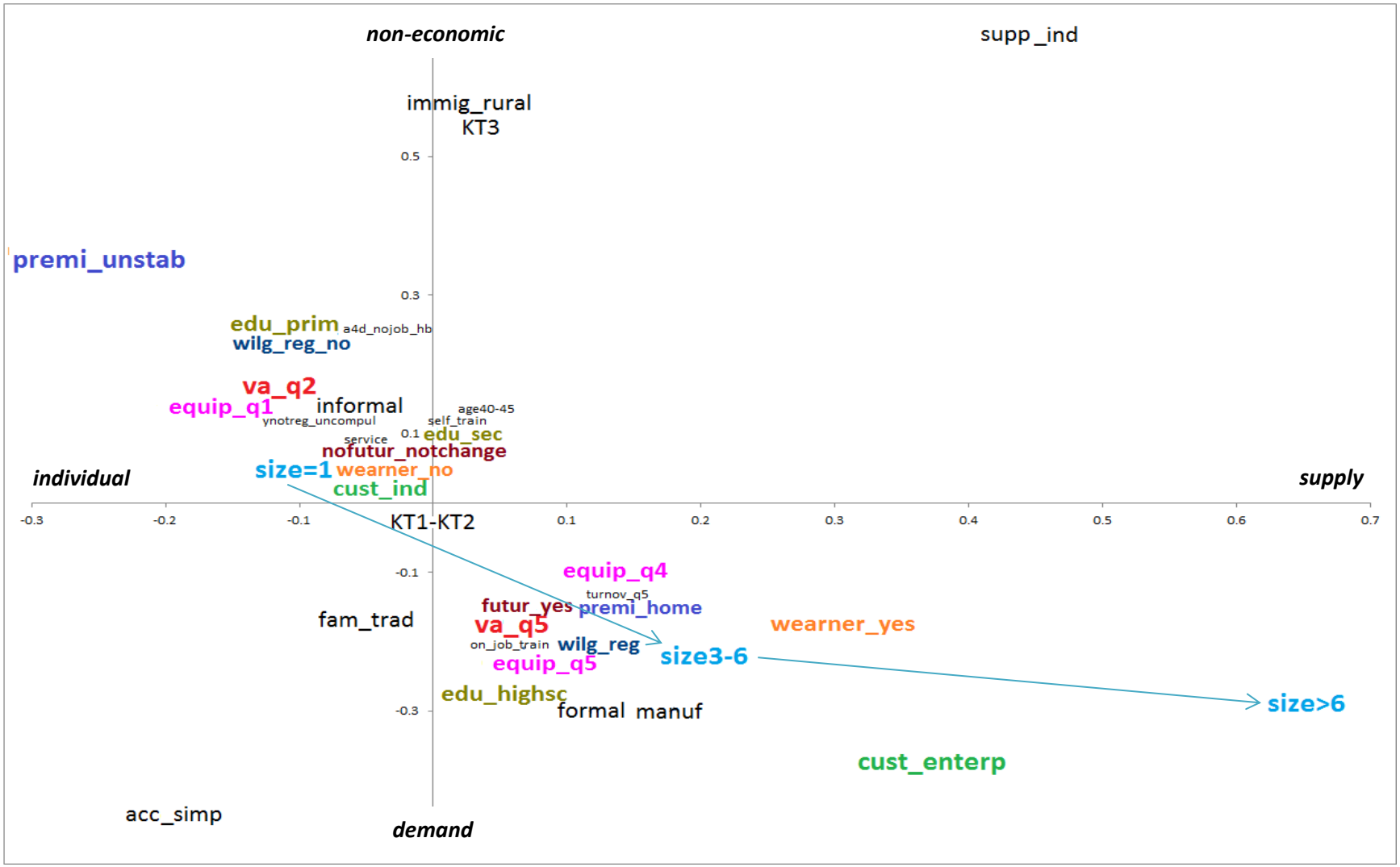
Source: HB&IS HCMC (2009); author' calculations. Note: The explicit meaning of the categorical variables is provided in Table 18 of Appendix.

Figure 5. The Space of Disadvantages Put Forward by HBs (First Factorial Plane of Correspondence Analysis)



Source: HB&IS HCMC (2009); author' calculations.

Figure 6. The Space of HB Characteristics (First Factorial Plane of Correspondence Analysis on Disadvantages)



Source: HB&IS HCMC (2009); author' calculations. Note: The explicit meaning of the categorical variables is provided in Table 18 of Appendix.

Combined corpus of advantages and disadvantages

For reference, the results of textual correspondence analysis on the combined corpus from the advantages and disadvantages of HBs are briefly presented. Although it is impossible to tell words that indicate advantages from words that capture disadvantages, the recorded words reflect what HB heads cared for, as well as what were important and meaningful to them. Since the combined corpus includes a higher number of distinct words than do separate corpuses, implying a higher dimensionality of textual data, the explanatory power of first principal axes significantly decreases compared to the previous analyses.[30] Yet, the factors remain fairly meaningful. **Table 10** shows the words that have the most contribution to the interpretation of the first two factorial axes.

Table 10. Selective Points with Extreme Contributions on the First Two Principal Axes: correspondence analysis on "Advantages" and "Disadvantages" of Informal HBs

Axis 1			Axis 2		
Words	**Translation**	**Coordinate**	**Words**	**Translation**	**Coordinate**
Dậy	get up	-22.027	Liên tục	continuously	-15.974
Sớm	early	-21.322	Thay đổi	change	-15.435
Thức	awake	-17.12	Dậy	get up	-13.067
Sáng	morning	-10.981	Sớm	early	-12.56
Cấm	ban	-9.307	Thức	awake	-10.11
Xe ba bánh	tricycle	-8.476	Nhân công	manpower	-9.786
Vất vả	strenous	-7.43	Tay nghề	skill	-8.012
Nhà nước	government	-6.372	Sáng	morning	-6.374
Chạy	run, drive	-3.378	Gia đình	family	-3.762
Đường	road, street	-3.092	Bệnh	disease	-3.335
Diện tích	surface area	2.28	Bụi	dust	3.131
Cạnh tranh	competition	2.289	Trời	weather	3.631
Doanh thu	revenue	2.62	Xe	vehicle	3.735
Địa điểm	location	3.02	Nắng	sunny	3.814
An ninh	security	3.231	Cúm	flu	3.873
Nhà	home	3.339	Dịch	epidemic	4.473
Tay nghề	skill	5.059	Dẹp	suppress	4.84
Nhân công	manpower	6.367	Lề	pavement	6.049
Liên tục	continuously	9.853	Mưa	rain	7.324
Thay đổi	change	9.983	Đường	road, street	9.057

Source: HB&IS HCMC (2009); author' calculations.

The first factorial axis seemingly differentiates *personal* and *entrepreneurial* perspectives. On the negative side of the axis, HBs tend to care about individual issues such as having to 'get up' 'early' in the 'morning' or the 'government' banning the 'tricycles' running on 'streets.' In contrast, those who are located on the positive side of this axis tend to put more professional topics from

[30] The five most variance-explained factors now contribute only 5.23 percent of the total variance of the corpus.

an entrepreneurial perspective, including but not limited to 'manpower,' 'location,' 'revenue,' 'competition' and 'surface area' (of the premise).

The second factorial axis distinguishes *internal* and mostly *manpower* from *external*, *environmental* matters. HBs positioned toward the negative side of the axis seem to care about their internal issues, mostly related to their owner (for examples, 'get up' 'early,' 'family') and the other workers ('manpower' 'continuously' 'change'). Whereas, HBs in the opposite side of the axis appear to pay attention to the surrounding environment which affects the operation of their business such as 'weather,' 'flu' 'epidemic' and 'dust.'

From the interpretation of these two factorial axes, it is possibly the case that HBs located toward the entrepreneurial and internal directions are *higher-end*, and vice versa. However, due to the combination of advantages and disadvantages, some mentioned terms remain controversial. For instance, 'streets' may signal a disadvantageous situation when it points to the precarious operating condition of street vendors; but probably at the same time refers to a professional shop located on a main street with many passengers, who become their potential customers. Therefore, the so-called 'continuum' of informal HBs may be hard to define or visualize in the first factorial space generated by the combined corpus.

To examine the aforementioned hypothesis, we project the illustrative categories on the first factorial plane as we did in **Figure 4** and **Figure 6**. Not surprisingly, the supplementary categories are distributed mostly around the origin of the factorial plane, implying little correlation between the textual factorial axes with the quantitative characteristics of HBs in this case.

4.3. Cluster analysis

Based on the dimensional space generated by correspondence analysis, cluster analysis is then applied to obtain a typology of informal HBs. Lebart et al. (2006) propose a test with the null hypothesis assuming that the average coordinate of a cluster on a factor is equal to zero, i.e., the profile of that cluster on this factor does not differ from the profile of the whole sample. If a cluster has a test absolute value greater than or equal to 1.65, the null hypothesis is rejected at five percent level of significance. The number of clusters can be controlled by cutting the dendrogram at a certain height. In this paper, the informal HBs are divided into five distinct clusters, based on either their advantages or their disadvantages.[31] The results for disadvantages seem to be more fruitful and implicative, thus being analyzed in detail in this section. For advantages, a summary of the results is given in **Table 20**, **Table 21** and **Table 22** (Appendix).

Table 11 shows that the profile of the clusters is significantly different from the global (i.e., their test absolute values are greater than 1.65), except the second cluster on the first factorial axis. The position of the clusters in the factorial space is visualized in **Figure 7**. Each cluster is characterized by their representative responses and then depicted by significant quantitative categories (**Table 12**). A similar v-test is computed in order to assess whether the percentage of

[31] For disadvantages, the number of clusters is initially determined equal to six. Then, there appears one HB which stands alone as a cluster. We consider this HB as an 'outlier' and finally, there are five remaining clusters whose results are reported here.

respondents in a class holding a modality is significantly different from that of the whole sample or not.

Table 11. Positions of Clusters in the Dimensional Space –Disadvantages of Informal HBs

Clusters		Coordinates		Test-values	
N°	Size	1	2	1	2
1	260	-0.39	-0.23	-12.6	-7.6
2	103	0.08	0.45	1.35	7.23
3	89	0.58	0.3	8.42	4.42
4	42	0.43	0.57	4.05	5.54
5	10	2.7	-3.63	12.05	-16.5

Source: HB&IS HCMC (2009); authors' calculations.

The first cluster covers *'the impervious HBs,'* who are more likely to declare that "There is no difficulty" obstructing their business. More than half of the HBs in this cluster have their premise at home, which appears to be a major advantage as suggested by the previous section. This is the most populous among five groups, with 260 HBs being classified in this cluster.

The second cluster comprises *'the optimistic HBs,'* who are more inclined to put forward their disadvantages in terms of prices, costs and profit. A substantial share of HBs in this group (40.78 percent) think that there is a future for a business like theirs. HBs in this cluster are more prone to achieve very high economic performance (the fifth quintile of value added).

The third cluster is made up of *'the desolate HBs,'* which includes the HBs who are more likely to face a lack of customers. They can be a tailor, who complains that people are increasingly buying cheap ready-made clothes in the marketplaces; or a motorbike taxi driver, who has less and less customers because his vehicle gets too old. Many more of them have a turnover of the lowest quintile (38.2 percent), compared to the average (26.73 percent).

The fourth cluster is characterized by *'the precarious HBs,'* who typically complain about the slow consumption of their goods due to bad weather (i.e., rain) and urban police expelling street vendors. This is consistent with their characteristic modalities. A majority of HBs in this group (76.19 percent) have unstable premises. These HBs have among the most precarious working conditions. Predictably, they are nearly twice as likely as the average to conceive that there is no future for their business, and to think of changing activity.

The fifth cluster covers *'the strenuous HBs,'* who generally complain that they have to get up very early in the morning (e.g., at two AM) to work. Interestingly, up to half of these HBs keep silent about their future prospect, i.e., neither upset nor optimistic. There are only ten HBs are classified into this cluster.[32]

[32] For this reason, this cluster is not illustrated in the factorial plane.

Table 12. Categorical Characteristics of the Clusters – Cluster Analysis on Disadvantages of Informal HBs

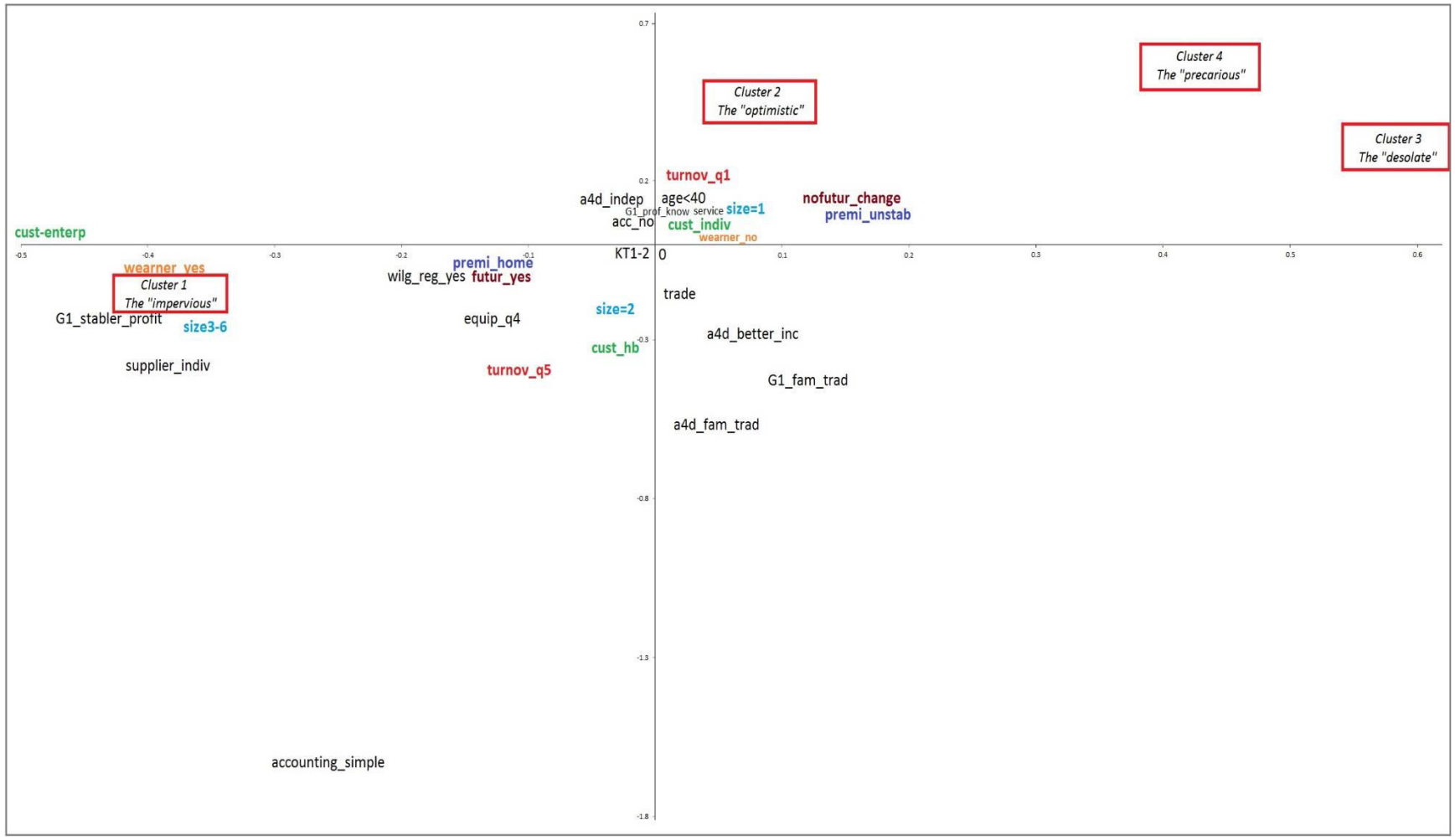
Characteristic categories	v.test	Percentages			Weight	Prob.
		cla/mod	mod/cla	global		
Cluster 1				51.49	260	
Individual supplier	2.59	83.33	5.77	3.56	18	0.0048
Premise at home	2.5	57.63	52.31	46.73	236	0.0062
Cluster 2				20.4	103	
Value added 5th quintile	2.68	37.78	16.5	8.91	45	0.0037
There is a future	2.64	28.19	40.78	29.5	149	0.0041
Cluster 3				17.62	89	
Turnover 1st quintile	2.5	25.19	38.2	26.73	135	0.0061
Cluster 4				8.32	42	
Unstable premise	4.98	16.33	76.19	38.81	196	0
No future, will change	2.45	15.31	35.71	19.41	98	0.0072
Cluster 5				1.98	10	
Missing value in question asking for "future"	2.87	50	20	0.79	4	0.0021

Source: HB&IS HCMC (2009); authors' calculations.

Note: "Cla/mod" indicates what share of all individuals with a certain category is found in a class. "Mod/cla" shows the percentage of all individuals in a cluster have a certain modality.

Overall, the results obtained by clustering analysis do provide further insights into the continuum of HBs suggested by correspondence analysis. This continuum is not as *'smooth'* as a usual continuous variable, since there is no absolute order of different clusters from the low-end to the high-end HBs. We could not tell, for example, which one among cluster 3 and cluster 4 is of lower end than the other. In fact, a HB has a variety of characteristics, some of them may be considered as high-end, whereas others are low-end. Therefore, it is hardly possible to have an absolute ranking of HBs in the continuum.

Figure 7. Cluster View of Informal HBs on the First Factorial Plane of Correspondence Analysis on Disadvantages



Source: HB&IS HCMC (2009); author' calculations. Note: The explicit meaning of the categorical variables is provided in Table 18 of Appendix.

4.4. Econometric models: How textual data can explain HB's economic performance?

This section aims to examine whether and how the textual data could help explain the economic performance of HBs in econometric models. The coordinates of HBs in the first two factorial axes generated from the textual correspondence analysis on the disadvantages of HBs are used as explanatory variables of economic outcomes. Recall that these axes reflect the perspective in which HB heads put forward and cope with their disadvantages, whether economic and entrepreneurial, or non-economic and personal; rather than those difficulties themselves. It could be said that the coordinates capture the level of business orientation or entrepreneurial skill of HB heads. For the first axis, the higher the coordinates, the higher level of entrepreneurial skill. For the second axis, the opposite holds. The coordinates are henceforth called *'entrepreneurial skill.'*[33] It is expected that the more entrepreneurial skill acquired by business owners, the better their business performs economically.

For comparison, we also examine the relationship between other measures of difficulties and economic performance of HBs. In principle, the problems declared by heads of HB are obstacles to the functioning of their business activity, thus having a close relationship with their economic achievement. In the HB&IS 2009 survey, a close-ended question (G4) asks if HBs face certain difficulties among a given list and their seriousness from 'no difficulty' (value=0) to 'very severe' (value=4). From this variable, we construct two measures of difficulties: (i) whether each type of difficulty exists, henceforth to be called *'type of difficulties*;' and (ii) sum of all these values, indicating *'degree of difficulty.'* The higher this indicator is, the more difficulties hindering HB's activity are perceived by their head.

The relationship between 'entrepreneurial skill' and economic performance of HBs

Table 13 presents the relationship between *'entrepreneurial skill,' 'type of difficulties*;' or *'degree of difficulty'* and economic performance of HBs, represented by monthly value added, with and without a set of controls. These controls include characteristics of HBs (informality, branch of activity, type of premise, age of HB), characteristics of HB owners (gender, education attainment, apprenticeship, migration status) and the motivation to create their business. Columns (1), (2), (3) show estimation results without conventional controls; and columns (5), (6), (7) present the full models with these controls. The benchmark model with only conventional quantitative explanatory variables is presented in column (4). In order to assess if *'entrepreneurial skill'* still matters when the type or intensity of difficulty are kept constant, we add *'type of difficulties'* and *'degree of difficulty* in columns (8) and (9). Finally, to determine the effect of entrepreneurial skill on productivity, factors of production are introduced in the regression equation in column (10).

To begin with, *'entrepreneurial skill'* appears to have a positive correlation with the economic performance of HBs. The first factor axis in model (1) of **Table 13** is statistically significant at ten percent level, indicating that HB heads who tend to put forward their disadvantages in an economic and business perspective would attain better economic outcome than those who emphasize their personal constraints. '*Type of difficulties'* is strongly correlated with monthly value

[33] *'Entrepreneurial capital'* is a broader concept than *'entrepreneurship capital,'* which means the capacities for economic agents to create new firms.

added in both models with and without controls. Among the pre-listed difficulties, *'lack of customers,' 'access to loan,' 'cash flow,' 'transportation'* and *'macroeconomic policies'* seem to matter the most to the performance of HBs. Surprisingly, while a *'lack of customers'* and difficult *'access to loan'* could reduce monthly value added, HB heads who state their problems in terms of *'cash flow,' 'transportation'* and *'macroeconomic policies'* on average achieve higher value added than those who do not (by 13.3 percent, 44.3 percent and 18.9 percent in the model without controls - column (2), and by 13.3 percent, 33.2 percent and 20.2 percent in the model with controls – column (6), respectively). These diverse results once again show that more professional and business-oriented difficulties tend to be declared by better-performed HBs.

Meanwhile, all other things equal, the degree or intensity of the difficulty stated actually does no harm to, and even benefits the economic performance of HBs. The coefficient of *'degree of difficulty'* is not significant in column (7) and positively significant at ten percent level in column (3). There are two possible explanations for this. First, the difficulties themselves do not impede the functioning of HBs if they can be overcome well. Second, owners of more successful HBs could probably better identify their obstacles and be more demanding and ambitious, thus mentioning more difficulties. For example, a PU owner who is aware of and cares about macroeconomic environment and policies such as inflation and exchange rate is likely to operate on a sufficiently large scale and under a professional management. In the qualitative HB&IS survey in 2009-2010, an interviewee mentioned the event of Vietnam joining the World Trade Organization and the fluctuation in the price of USD as disadvantageous to his repairing service, and he had two different professional premises in the same street (Nguyen Van Xuan, selling and repairing air conditioner).

It is further confirmed that even when types and degree of difficulties are controlled for, *'entrepreneurial skill'* remains beneficial to the economic performance of HBs. The first factorial axes in models (8) and (9) are positively significant at ten percent level. In other words, confronting the same amount or types of difficulty, those who can face them in a more entrepreneurial regard could possibly overcome their problems more easily and achieve a better economic outcome than those who perceive them as personal issues. Recall that our observation units are called *'household businesses,'* implying that their business is more or less attached to their household with personal and familial constraints as well as thinking. The more HB owners can detach from these constraints and thinking, the more likely they can approach professional enterprises.

To investigate how *'entrepreneurial skill'* could improve the economic performance of HBs, we first examine the relationship between entrepreneurial skill and HB's productivity. Column (10) of **Table 13** shows that using the same quantity of production factors (i.e., labor and equipment), on average, HBs with different amount of entrepreneurial skill would attain similar levels of value added, *ceteris paribus*. The fact that entrepreneurial skill enhances economic performance without promoting productivity suggests that this kind of skill probably influences economic performance through other transmission channels. **Table 14** and **Table 15** examine the correlation between HB head's entrepreneurial skill and their employment of labor and physical capital.

As expected, entrepreneurial skill is positively correlated with the total size and investment in equipment of the PUs. Both axes are highly significant in various specifications. This shows that as HBs become more business-oriented, their economic performance is improved through

enlarging scale of production rather than through upgrading productivity. Although productivity enhancement could be seen as a popular advantage of business expansion, it should be reminded that the studied units are HBs, who, according to the law, *"may employ only up to ten employees, shall not have a seal, and shall be liable for its business activities to the full extent of its assets."* They are not yet professional enterprises in essence. Therefore, it is likely that the concept of *'entrepreneurial skill'* in HBs' context could hardly lead to technological breakthroughs. Instead, a greater amount of *'entrepreneurial skill'* or business orientation allows HBs to increase production factors and hence, final output.

Whereas, the degree of difficulty has no relationship with investment in equipment and positive correlation with total size of HBs. Looking at specific types of difficulty, mentioning *'macroeconomic policies'* appears to be a signal of a larger amount of both production factors. In terms of labor size, *'transportation'* problem is more frequently stated by larger HBs. Meanwhile, *'too much competition'* and *'lack of machines and equipment'* concerns turn out to be associated with a higher level of investment in equipment. This example illustrates the specificity of subjective well-being: sometimes, complaints indicate ambition rather than objective difficulty.

In brief, econometric results have shown that what matters to the economic performance of HBs may be not the amount and degree of difficulties they face, but the perspective in which they perceive and deal with these problems. The latter could be considered as their *'entrepreneurial skill'* or business orientation. *'Entrepreneurial skill'* seems to improve HBs' performance by allowing them to expand their factors of production rather than enhancing their productivity.

Table 13. Determinants of Economic Performance of HBs

	Value added (monthly, in logarithm)									
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Entrepreneurial skill										
First factor axis	0.105*				0.046			0.094*	0.095*	-0.013
	(0.054)				(0.048)			(0.054)	(0.054)	(0.044)
Second factor axis	-0.064				0.057			-0.043	-0.065	0.022
	(0.051)				(0.048)			(0.051)	(0.051)	(0.042)
Close-ended question: Difficulties										
Supply of raw materials		0.105				0.138*		0.102		
(quantity or quality)		(0.084)				(0.076)		(0.084)		
Lack of customers		-0.199***				-0.140***		-0.190***		
		(0.051)				(0.046)		(0.051)		
Too much competition		0.060				-0.032		0.056		
		(0.044)				(0.041)		(0.045)		
Access to loan		-0.196***				-0.106*		-0.204***		
		(0.062)				(0.056)		(0.063)		
Cash flow		0.133**				0.133**		0.127**		
		(0.062)				(0.056)		(0.062)		
Lack of space and appropriate premise		-0.036				0.024		-0.038		
		(0.059)				(0.053)		(0.059)		
Lack of machines and equipment		-0.058				-0.149**		-0.068		
		(0.074)				(0.069)		(0.074)		
Transportation		0.443***				0.332***		0.446***		
		(0.102)				(0.093)		(0.102)		
Access to land		0.150				0.067		0.137		
		(0.101)				(0.090)		(0.101)		
Macroeconomic policy		0.189***				0.202***		0.187***		
(inflation, exchange rate)		(0.069)				(0.062)		(0.069)		

Crime, theft and disorder	0.009			-0.058		0.023	
	(0.091)			(0.082)		(0.091)	
Other difficulties	-0.015			-0.060		-0.025	
	(0.092)			(0.082)		(0.092)	
Degree of difficulty		0.020*			0.013		0.018*
		(0.010)			(0.009)		(0.011)
Characteristics of HB							
Informal			-0.792***	-0.801***	-0.766***	-0.781***	
			(0.092)	(0.093)	(0.092)	(0.093)	
Branch of activity (Reference: Manufacturing)							
Trade			-0.359***	-0.364***	-0.361***	-0.371***	
			(0.111)	(0.112)	(0.112)	(0.112)	
Service			-0.155	-0.162	-0.126	-0.159	
			(0.105)	(0.105)	(0.103)	(0.105)	
Type of premise (Reference: Precarious)							
Home			-0.141	-0.141	-0.085	-0.139	
			(0.098)	(0.099)	(0.097)	(0.098)	
Professional			0.185	0.181	0.197*	0.191	
			(0.116)	(0.117)	(0.114)	(0.116)	
Age of HB			0.005	0.005	0.001	0.005	
			(0.005)	(0.005)	(0.005)	(0.005)	
Characteristics of HB head							
Female			-0.247***	-0.248***	-0.250***	-0.251***	
			(0.074)	(0.074)	(0.074)	(0.074)	
Years of schooling			0.026**	0.026**	0.025**	0.025**	
			(0.011)	(0.011)	(0.011)	(0.011)	
Apprenticeship (Reference: Technical school)							
In enterprise or HB			0.213	0.208	0.033	0.225	
			(0.176)	(0.176)	(0.176)	(0.176)	
Alone, by practice			0.134	0.120	-0.040	0.140	

			(0.159)	(0.159)	(0.159)	(0.159)			
Others			-0.049	-0.043	-0.194	-0.032			
			(0.192)	(0.192)	(0.191)	(0.192)			
Migrant			0.264	0.250	0.275*	0.265			
			(0.162)	(0.162)	(0.162)	(0.162)			
Motivation to create HB (Reference: Did not find salaried work in enterprise)									
Did not find salaried work in HB			-0.303**	-0.308**	-0.316**	-0.300**			
			(0.151)	(0.151)	(0.149)	(0.151)			
To get a better income			0.239*	0.239*	0.174	0.248*			
			(0.130)	(0.130)	(0.128)	(0.130)			
To be independent			0.296***	0.289***	0.232**	0.300***			
			(0.107)	(0.107)	(0.106)	(0.107)			
By family tradition			0.136	0.133	0.166	0.142			
			(0.153)	(0.153)	(0.151)	(0.153)			
Other motivations			-0.209	-0.224	-0.382**	-0.205			
			(0.177)	(0.177)	(0.176)	(0.177)			
Factors of production									
Total size									0.329***
									(0.020)
Investment in equipment									0.053***
									(0.013)
Constant	8.250***	8.215***	8.612***	8.633***	8.808***	8.566***	8.220***	8.192***	7.107***
	(0.040)	(0.052)	(0.268)	(0.268)	(0.268)	(0.270)	(0.052)	(0.053)	(0.123)
Observations	740	740	729	729	729	729	740	740	688
R-squared	0.007	0.081	0.257	0.259	0.309	0.259	0.086	0.011	0.356

*Source: HB&IS HCMC (2009); author' calculations. Standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1.*

Table 14. Determinants of Total Size of HBs

	Total size of HB								
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Entrepreneurial skill									
First factor axis	0.276***				0.164**			0.270***	0.257***
	(0.085)				(0.080)			(0.086)	(0.085)
Second factor axis	-0.160**				0.018			-0.125	-0.161**
	(0.081)				(0.079)			(0.082)	(0.081)
Close-ended question: Difficulties									
Supply of raw materials		0.080				0.081		0.072	
(quantity or quality)		(0.136)				(0.129)		(0.136)	
Lack of customers		-0.145*				-0.093		-0.119	
		(0.082)				(0.078)		(0.081)	
Too much competition		0.077				0.030		0.067	
		(0.072)				(0.070)		(0.072)	
Access to loan		-0.157				-0.061		-0.179*	
		(0.101)				(0.094)		(0.101)	
Cash flow		0.113				0.166*		0.096	
		(0.100)				(0.095)		(0.100)	
Lack of space and appropriate premise		-0.138				-0.054		-0.141	
		(0.095)				(0.090)		(0.095)	
Lack of machines and equipment		0.011				-0.173		-0.019	
		(0.119)				(0.117)		(0.118)	
Transportation		0.665***				0.523***		0.674***	
		(0.165)				(0.157)		(0.164)	
Access to land		0.223				0.112		0.187	
		(0.163)				(0.152)		(0.162)	
Macroeconomic policy		0.198*				0.230**		0.195*	
(inflation, exchange rate)		(0.111)				(0.106)		(0.110)	

Crime, theft and disorder	-0.002				-0.123		0.039	
	(0.146)				(0.139)		(0.146)	
Other difficulties	0.140				0.114		0.110	
	(0.148)				(0.139)		(0.147)	
Degree of difficulty		0.040**				0.039**		0.035**
		(0.017)				(0.016)		(0.017)
Characteristics of HB								
Informal			-0.949***	-0.941***	-0.877***	-0.916***		
			(0.154)	(0.155)	(0.156)	(0.154)		
Branch of activity (Reference: Manufacturing)								
Trade			-1.085***	-1.066***	-1.142***	-1.120***		
			(0.186)	(0.187)	(0.190)	(0.186)		
Service			-0.733***	-0.720***	-0.712***	-0.743***		
			(0.175)	(0.176)	(0.175)	(0.175)		
Type of premise (Reference: Precarious)								
Home			-0.015	-0.050	0.030	-0.010		
			(0.163)	(0.165)	(0.164)	(0.163)		
Professional			0.324*	0.289	0.337*	0.343*		
			(0.193)	(0.194)	(0.192)	(0.193)		
Age of HB			-0.004	-0.004	-0.009	-0.005		
			(0.008)	(0.008)	(0.008)	(0.008)		
Characteristics of HB head								
Female			-0.354***	-0.352***	-0.338***	-0.364***		
			(0.124)	(0.124)	(0.126)	(0.124)		
Years of schooling			0.032*	0.032*	0.028	0.029		
			(0.018)	(0.018)	(0.018)	(0.018)		
Apprenticeship (Reference: Technical school)								
In enterprise or HB			0.427	0.407	0.249	0.462		
			(0.294)	(0.294)	(0.298)	(0.294)		
Alone, by practice			0.335	0.305	0.139	0.354		

				(0.265)	(0.265)	(0.270)	(0.264)		
Others				0.328	0.316	0.167	0.380		
				(0.320)	(0.320)	(0.324)	(0.320)		
Migrant				0.670**	0.669**	0.728***	0.675**		
				(0.267)	(0.267)	(0.271)	(0.266)		
Motivation to create HB (Reference: Did not find salaried work in enterprise)									
Did not find salaried work in HB				-0.020	-0.029	-0.000	-0.013		
				(0.252)	(0.252)	(0.252)	(0.251)		
To get a better income				0.408*	0.423*	0.367*	0.434**		
				(0.217)	(0.217)	(0.217)	(0.216)		
To be independent				0.468***	0.462***	0.431**	0.480***		
				(0.178)	(0.178)	(0.180)	(0.178)		
By family tradition				0.499*	0.509**	0.573**	0.516**		
				(0.256)	(0.256)	(0.257)	(0.255)		
Other motivations				0.143	0.120	-0.041	0.155		
				(0.296)	(0.296)	(0.298)	(0.295)		
Constant	1.919***	1.823***	1.790***	2.496***	2.527***	2.610***	2.358***	1.838***	1.806***
	(0.064)	(0.084)	(0.083)	(0.447)	(0.447)	(0.454)	(0.449)	(0.084)	(0.083)
Observations	741	741	741	730	730	730	730	741	741
R-squared	0.019	0.051	0.008	0.192	0.197	0.226	0.199	0.066	0.025

*Source: HBe&IS HCMC (2009); author' calculations. Standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1.*

Table 15. Determinants of Investment in Equipment of HBs

	Investment in equipment in the last 12 months								
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Entrepreneurial skill									
First factor axis	0.392***				0.083			0.432***	0.372***
	(0.138)				(0.118)			(0.140)	(0.140)
Second factor axis	-0.501***				-0.041			-0.400***	-0.504***
	(0.132)				(0.117)			(0.133)	(0.132)
Close-ended question: Difficulties									
Supply of raw materials		-0.135				-0.094		-0.176	
(quantity or quality)		(0.226)				(0.194)		(0.224)	
Lack of customers		-0.307**				-0.349***		-0.271**	
		(0.135)				(0.117)		(0.134)	
Too much competition		0.354***				0.248**		0.319***	
		(0.118)				(0.104)		(0.118)	
Access to loan		-0.518***				-0.365**		-0.531***	
		(0.167)				(0.142)		(0.168)	
Cash flow		0.189				0.116		0.176	
		(0.167)				(0.144)		(0.165)	
Lack of space and appropriate premise		-0.397**				-0.097		-0.392**	
		(0.159)				(0.136)		(0.158)	
Lack of machines and equipment		0.421**				0.042		0.356*	
		(0.193)				(0.173)		(0.191)	
Transportation		0.344				0.320		0.361	
		(0.282)				(0.243)		(0.278)	
Access to land		0.301				0.281		0.231	
		(0.270)				(0.229)		(0.267)	
Macroeconomic policy		0.320*				0.354**		0.323*	
(inflation, exchange rate)		(0.187)				(0.161)		(0.184)	

Crime, theft and disorder	0.280				0.157		0.360	
	(0.244)				(0.209)		(0.242)	
Other difficulties	0.422*				0.261		0.357	
	(0.247)				(0.210)		(0.245)	
Degree of difficulty		0.038				0.027		0.031
		(0.028)				(0.024)		(0.028)
Characteristics of HB								
Informal			-0.883***	-0.869***	-0.684***	-0.861***		
			(0.230)	(0.232)	(0.232)	(0.231)		
Branch of activity (Reference: Manufacturing)								
Trade			-0.021	-0.003	-0.037	-0.042		
			(0.280)	(0.282)	(0.286)	(0.281)		
Service			0.342	0.357	0.311	0.336		
			(0.264)	(0.265)	(0.263)	(0.264)		
Type of premise (Reference: Precarious)								
Home			2.825***	2.799***	2.883***	2.825***		
			(0.249)	(0.251)	(0.249)	(0.248)		
Professional			1.961***	1.938***	1.960***	1.976***		
			(0.294)	(0.295)	(0.293)	(0.294)		
Age of HB			-0.016	-0.016	-0.024**	-0.017		
			(0.012)	(0.012)	(0.012)	(0.012)		
Characteristics of HB head								
Female			-0.390**	-0.389**	-0.391**	-0.400**		
			(0.187)	(0.187)	(0.189)	(0.187)		
Years of schooling			0.040	0.039	0.030	0.039		
			(0.028)	(0.028)	(0.027)	(0.028)		
Apprenticeship (Reference: Technical school)								
In enterprise or HB			0.409	0.396	0.333	0.437		
			(0.433)	(0.434)	(0.439)	(0.433)		
Alone, by practice			-0.504	-0.515	-0.511	-0.489		

				(0.391)	(0.393)	(0.398)	(0.391)		
Others				-0.904*	-0.923*	-1.012**	-0.872*		
				(0.480)	(0.481)	(0.485)	(0.481)		
Migrant				-0.701*	-0.690*	-0.771*	-0.702*		
				(0.400)	(0.401)	(0.407)	(0.400)		
Motivation to create HB (Reference: Did not find salaried work in enterprise)									
Did not find salaried work in HB				0.002	-0.003	0.153	0.010		
				(0.393)	(0.394)	(0.393)	(0.393)		
To get a better income				0.541	0.551*	0.631*	0.565*		
				(0.329)	(0.329)	(0.329)	(0.329)		
To be independent				0.667**	0.667**	0.687**	0.677**		
				(0.272)	(0.273)	(0.275)	(0.272)		
By family tradition				0.726*	0.733*	0.910**	0.739*		
				(0.389)	(0.389)	(0.389)	(0.389)		
Other motivations				1.102**	1.093**	0.936**	1.107**		
				(0.456)	(0.457)	(0.460)	(0.456)		
Constant	9.842***	9.693***	9.719***	8.369***	8.377***	8.281***	8.267***	9.726***	9.743***
	(0.106)	(0.142)	(0.141)	(0.666)	(0.668)	(0.677)	(0.672)	(0.141)	(0.140)
Observations	689	689	689	678	678	678	678	689	689
R-squared	0.031	0.062	0.003	0.337	0.338	0.367	0.338	0.087	0.032

*Source: HB&IS HCMC (2009); author' calculations. Standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1.*

Factorial axes as a measure of entrepreneurial skill

To assess the validity of factorial axes as measures of *'entrepreneurial skill,'* we use another proxy of business orientation in the HB&IS questionnaire. In fact, the HB&IS questionnaire 2009 includes a question capturing loan usage behavior of HBs, letting HBs to choose among different productive and non-productive purposes (G7). Simple regression shows that the higher level of entrepreneurial skill (indicated by larger value of the first factorial axis), the more probable HBs use their loan for a productive and business-related purpose. **Table 16** also presents average values of coordinates on the first factorial axis corresponding to each loan use decision. It points out that HBs who allocate their loan for non-productive, non-business expenditure have significantly lower entrepreneurial skill than those who use their credit to improve premise and site (-0.128 versus 0.157, five percent level of significance). This statistics supports our employment of factorial axes generated from the textual correspondence analysis as a measure of entrepreneurial skill owned by HB heads, thus affirming our previous findings.

Table 16. 'Entrepreneurial capital' and loan usage behavior

'Entrepreneurial capital'	Mean	Std. Err.	95% Confidence Interval	
Increase stock of raw materials	-0.052	0.061	-0.172	0.068
Improve premise, site	0.157	0.062	0.036	0.277
Improve machines, furniture, tools	0.056	0.068	-0.077	0.189
Hire workers	0.348	0.283	-0.207	0.903
Open another establishment of the same activity	0.002	0.118	-0.230	0.233
Open another establishment of different activity	0.126	0.154	-0.177	0.428
Non-productive, non-business expenditure	-0.128	0.046	-0.219	-0.037
Other	-0.023	0.082	-0.183	0.137

Source: HB&IS HCMC (2009); author' calculations.

The fact that entrepreneurial skill can be to some extent reflected by types of difficulty and loan usage behavior suggests the trade-off between using open-ended questions and close-ended ones. On the one hand, the open-ended questions may be preferred to conventional close-ended ones in the sense that the former might reveal richer information and limit the pre-conceived ideas of researchers in designing the questionnaire relative to the latter. On the other hand, while given the freedom to write, the interviewees may forget or intentionally ignore certain aspects (which are mentioned in the quantitative question). Moreover, during the process of transforming texts into factorial axes, certain information might be lost. Indeed, adding the factorial axes to the conventional set of determinants of production factors (i.e., quantitative variables representing HB's and HB owner's characteristics) seems not to improve the overall explanatory power of the models (**Table 13**, **Table 14** and **Table 15**). Meanwhile, adding *'type of difficulties'* as independent variables could raise the R-squared by around three percent.

However, this does not necessarily mean that the responses to open-ended questions reveal less information concerning economic performance and functioning of HBs than do close-ended questions. In fact, during the transformation process from text corpus to lexical table and from lexical table (frequency matrix) to the factorial axes, a certain amount of information has been dropped. The original advantages and disadvantages revealed by HBs are rich, vivid and embedded

with emotions and personal attitudes. Meanwhile, the textual analysis method only focuses on the juxtaposition or the co-occurrence of terms as well as repeated segments (Rakotomanana & Rakotoasimbola, 2015). The coordinates generated with the software program have been constructed based on the frequency of words, ignoring their situational semantic meanings and the semantic order of phrases or expressions. Therefore, the created coordinates, differently speaking, the quantified form of qualitative data, might not take the most advantage of text corpus.

4.5. Robustness check

This section aims at examining whether the *'entrepreneurial skill* is still captured by the factorial axes and the continuum of HBs in the informal sector remains evident in response to some modifications in data construction. Several adjustments could be implemented. First, as far as textual data cleaning as concerned, we may: (i) modify the list of 'stop-words' and the list of synonyms to be merged; and (ii) change the frequency threshold of words to be kept. Second, some categorical variables could be re-grouped or added. For examples, indicators of economic performance (e.g., turnover) can be divided into quartiles instead of quintiles, and the informality trajectory of HBs between 2007 and 2009 could serve as an additional supplementary variable. Third, certain categorical variable(s) representing personal viewpoints of HB heads could be used in correspondence analysis as if they were textual ones. Differently speaking, instead of playing the role of supplementary variables, these categories are 'textualized' so that they become active variables in textual correspondence analysis, together with the open-ended question(s). In principle, while the first and third interventions could directly affect the formation of the factorial axes, the second one only changes the illustration of the factorial spaces and description of clusters.

We continue with the analyses on the disadvantages of informal HBs presented in sections 4.2 and 4.3. First, after deleting more 'stop-words,' merging more words with comparable meanings and imposing a higher frequency threshold of kept words (i.e., ten instead of two), the total variance explained by the first five factors has increased to 16.26 percent (compared to 6.52 percent in section 4.2). This is because the aforementioned steps have significantly reduced the dimensionality of the corpus.

Then, the motivation of creating the HB (variable A4d) is transformed into words and becomes the second open-ended question in the textual correspondence analysis in DTM-Vic.[34] We choose this variable because it is directly related to the disadvantages of HBs, revealing their 'starting point': whether it is a choice or a constraint. It is expected that the constraint modalities appear in the *'low-end'* side of the axes and reversely. The modalities of variable A4d are recoded as 'words' in STATA (i.e., letters are linked together without any space between them) before being imported to DTM-Vic. Each numerical modality is replaced by such a 'word.'[35] We also tried using other variables such as G1 (The main reason to choose this business activity) and G31e (The main

[34] By doing so, the total variance explained by the first five factors has slightly decreased to 14.66 percent. This is because the number of distinct words, i.e. dimensions, in the corpus has increased a little.

[35] The generated 'words' and their corresponding meanings are: *'notfindworkenterp'* (did not find work in an enterprise), *'notfindworkHB'* (did not find work in a HB), *'betterincome'* (to get higher income), *'independent'* (to be independent), *'a4dfamtrad'* (by family tradition), *'a4dother'* (other reasons).

benefit of registering business), but including too many categories as active variables in the textual correspondence analysis actually biases the axes towards these generated 'words.' To be detailed, these 'words' become the most frequent ones in the corpus, thus contributing the most to the interpretation of the factorial axes, leaving the 'real' words poorly represented. For this reason, we decide to textualize only variable A4d.

Table 17 and **Figure 8** present the results of textual correspondence analysis on the disadvantages and motivations of informal HBs with the augmented list of stop-words and synonyms and the new (i.e., higher) frequency threshold. In general, we still obtain meaningful factors and the projection of illustrative variables seems to support the interpretation of the factorial axes. *'Entrepreneurial skill* remains reflected through the factorial axes and a continuum of HBs in the informal sector stays evident.

The first factorial axis distinguishes *survivor* from *ambitious* HBs. Those who are located toward the negative side of the axis are more likely to operate their business for survival objective. Their owners tend to set up business due to a constraint (i.e., did not find work elsewhere) and put forward external difficulties such as 'street,' 'pavement,' 'dust,' 'rain,' 'petrol' in a seemingly passive manner. Meanwhile, those who stay toward the positive side of this axis tend to mention professional concerns (e.g., 'capital,' 'premise,' 'production') and express their ambitions (e.g., 'enlarge' the business, motivations: better income, family tradition). This is in line with the study on craft village enterprises (which are mainly HBs) by Vu et al. (2010), which shows that family tradition plays a positive role in improving business scale and product quality.

The second factorial axis separates out *economic* and *non-economic* problems. On the negative side, factors such as 'price,' 'profit,' 'income' and 'rent' occupy the HB owners the most. On the positive side, HB heads are more likely to worry about non-economic factors such as 'pavement,' 'rain,' 'age,' and so on. In fact, these so-called 'non-economic' factors could possibly be considered as 'economic' in the sense that they represent the operating environment or endowment of the interviewees. The differences between the two groups of disadvantages might suggest that the two HB groups are in different stages of development. For the latter group, exogenous factors which are unable to be intentionally improved by the HBs themselves could impede their operation. This indicates the vulnerability of these HBs relative to the former group.

Table 17. Selective Points with Extreme Contributions on the First Two Principal Axes: Correspondence Analysis on "Disadvantages" and Motivations of Informal HBs

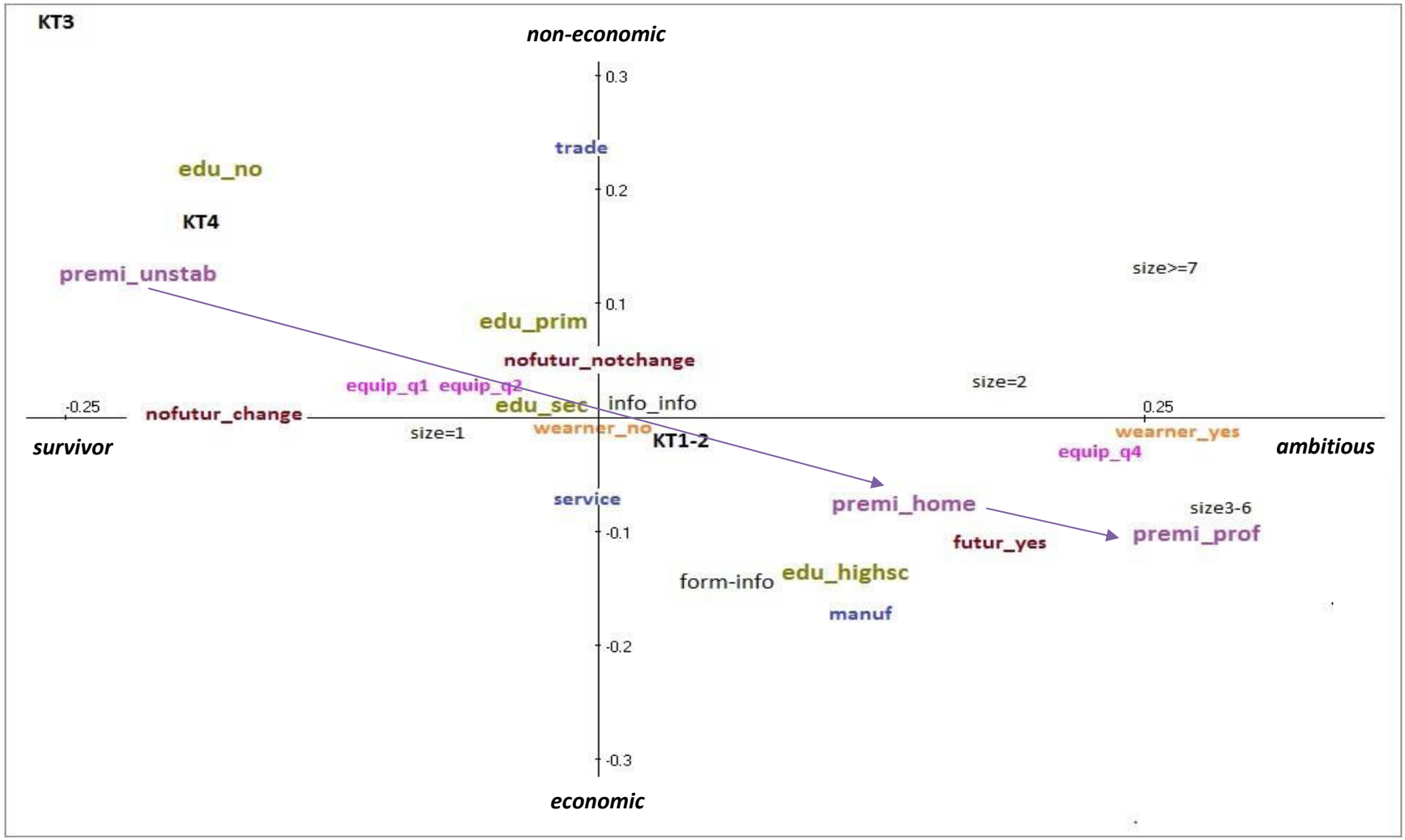
Axis 1			Axis 2		
Words	**Translation**	**Coordinate**	**Words**	**Translation**	**Coordinate**
Đường	street, road	-11.476	Giá	price	-11.837
Lề	pavement	-10.353	Khách	customer	-5.443
Lòng	middle (of the road)	-9.296	Vật giá	price of inputs	-4.201
Xe	vehicle	-5.914	Xăng	petrol	-3.791
Bụi	dust	-5.542	Lợi nhuận	profit	-3.564
Mưa	rain	-5.355	Lượng	quantity	-3.034
Khách	customer	-4.248	Thu nhập	income	-2.814
Động lực: Không tìm được	Not find work in HB (A4d)	-4.218	Nước	water	-2.799

việc ở hộ sản xuất kinh doanh					
Chậm	slow	-3.547	Hàng	product, goods	-2.176
Xăng	petrol	-3.483	Thuê	rent	-2.057
Giảm	reduce	-2.858	Động lực: thu nhập cao hơn	Better income (A4d)	-1.785
Thu nhập	income	-2.624	Xe	vehicle	-1.734
Ổn định	stable	-2.184	Dân	population, inhabitant	-1.709
Thời gian	time	-1.989	Cạnh tranh	competition	-1.202
Động lực: Không tìm được việc ở doanh nghiệp	Not find work in enterprise (A4d)	-1.259	Ngành nghề	activity, branch	-1.187
Cạnh tranh	competition	1.073	Động lực: Truyền thống gia đình	Family tradition (A4d)	1.306
Địa điểm	location	2.35	Bụi	dust	1.507
Lợi nhhuận	profit	2.611	Cố định	fixed	2.104
Tiền	money	2.775	Ế	not sellable	2.285
Tuổi	age	3.137	Chậm	slow	2.478
Nhà nước	government	3.746	Chỗ	place	2.527
Thuê	rent	3.953	Địa điểm	location	3.084
Động lực: thu nhập cao hơn	Better income (A4d)	4.045	Vốn	capital	3.212
Động lực: Truyền thống gia đình	Family tradition (A4d)	4.649	Nhà nước	government	4.6
Sản xuất	production, manufacturing	4.875	Tuổi	age	4.804
Mở rộng	enlarge	5.675	Cấm	to ban	5.278
Vay	borrow	6.216	Mưa	rain	6.144
Mặt bằng	premise	6.263	Đường	road, street	7.477
Vốn	capital	7.715	Lòng	middle of the road	7.581
Kinh doanh	business	8.992	Lề	pavement	8.808

Source: HB&IS HCMC (2009); author' calculations.

Figure 8 visualizes a continuum of HBs from the top left to the bottom right of the first factorial plane. On the one hand, lower-end HBs are generally more precarious (unstable premise, temporary residence registration) and poorer (first and second quartiles of investment in equipment). Their heads have lower levels of education (no education or primary school) and keep a pessimistic attitude toward their business future. These HBs tend to operate to sustain a living and be vulnerable to exogenous problems. On the other hand, the higher-end HBs exhibit a larger economic activity (highest quartile of investment in equipment) and enjoy more stable working conditions (premise at home or professional premise). Their owners have a longer educational attainment and are more optimistic about their business future. The disadvantages put forward by these HBs seem to reflect their business ambitions.

Figure 8. The Space of Informal HB Characteristics (First Factorial Plane of Correspondence Analysis on Disadvantages and Motivation)



Source: HB&IS HCMC (2009); author' calculations.

5. Conclusion

Informal sector plays a vital role in Vietnam's economy as an important source of living for the urban population, especially the poor. Yet, this sector has been understudied and overlooked by the State. Earlier studies have found that it is more complicated than simply a residual sector as suggested by the dualist school. Therefore, to define assistance policies targeting at this sector, the government needs to understand the sector's systematic features and diversity, which is also the aim of this research work. Based on a unique *quali-quanti* analytical approach, this paper has highlighted the heterogeneity of the informal sector in Vietnam.

This research is marked with several originalities. First, the analyses are based on representative textual data, which measure the direct utility of HBs in an open and unrestricted manner. The essence of this paper is derived from two open-ended questions showing advantages and disadvantages during the operation of HBs. This text corpus provides copious and unique information which is hardly seen in conventional quantitative data. It reveals personal experience, attitude as well as ambition of the respondents. Second, this is the first economic paper that handles the linguistic problems associated with Vietnamese texts, for which factor analysis can be applied. Last but not least, exploratory analytical methods (i.e, correspondence analysis and clustering technique) allow to understand the underlying structure of these data and limit the pre-conceived ideas of researchers. The textual data are 'quantified' into principal axes representing *'entrepreneurial skill,'* which can explain the variation in the economic performance and functioning of HBs in econometric models. To sum up, a combination of qualitative and quantitative data as well as methodologies has brought original results to this research work.

Textual correspondence analysis shows that HBs in the informal sector are located along a continuum in a multi-dimensional space of their disadvantages *or* advantages. HBs at the lower end of the continuum are generally poorer, smaller, more precarious and less integrated in the economy. Their owners are less educated and more pessimistic about their business future. These HBs often lower their ambition and choose less demanding work in order to achieve job satisfaction. Yet, low-end HBs still put forwards their constraints in a personal and passive perspective. Reversely, the higher-end HBs operate on a larger scale, have better economic performance, more stable working conditions and are more incorporated in the economy. Their owners have higher educational attainment and are more optimistic about their business future. These HBs tend to stress their difficulties in an entrepreneurial perspective, for that they can grasp opportunities and actively improve their situation. Furthermore, a cluster analysis is conducted on the disadvantages faced by HBs, providing a typology with five categories of informal HBs, namely: *'the impervious,' 'the optimistic,' 'the desolate,' 'the precarious'* and *'the strenuous.'*

Econometric analyses show that the heterogeneity in the economic performance of HBs could be explained by their different levels of *'entrepreneurial skill,'* captured by HBs' coordinates on principal factors generated from textual correspondence analysis on disadvantages. We find that what matters to the economic performance of HBs may be not the amount and degree of difficulties they confront, but the perspective in which they identify and cope with these problems. The more *'entrepreneurial skill'* or business orientation that a HB head expresses, the more likely their business achieves a high economic outcome. *'Entrepreneurial skill'* seems to improve HBs'

performance by allowing them to expand their scale of production rather than by enhancing their productivity.

In terms of policy implication, the obtained results directly suggest that targeted policies should take into account the heterogeneity of the informal sector in Vietnam. A 'one size fits all' scheme would not be appropriate, as there is not one single reason for working in this sector, and different categories of informal HBs experience different kinds of difficulties as well as advantages and have their own future perspective. Moving downward from the *high-end* to *low-end*, government first of all should adopt anti-poverty policy for the households, then it is more on infrastructure (e.g., roadways) and finally 'professional' policies ought to be exercised on macro business environment in order to, for example, confine inflation and so on. Likewise, the identification of several clusters of HBs calls for diverse measures which can encounter their specific types of disadvantages. A policy package may include, for examples, premise stabilization for *'precarious HBs,'* vocational training for *'desolate HBs'* and inflation control for *'optimistic HBs.'* Moreover, training courses which enhance the business and entrepreneurial skills of HBs could possibly help them improve their economic performance.

This research could be improved in several ways. First, the codification process is performed manually, which would be impractical for larger data sets. This calls for the development of statistical packages that incorporate cutting-edge linguistics' and informatics' achievements in order to properly treat Vietnamese language in a systematic way. Second, DTM-Vic does not allow to use both textual and quantitative data as active variables simultaneously, but we could do so with another software program. Third, one could take further advantage of the panel dimension of the HB&IS data to investigate the dynamics of the informal sector. Last but not least, the generated measure of HBs' *'entrepreneurial skill'* could be further exploited in other research subjects in economics and management study.

Appendix

Table 18. Categorical Variables Used in Correspondence Analysis

N°	Description	Modalities	Abbreviation
		Characteristics of HB	
1	Branch of activity	Manufacturing	*manuf*
		Trade	*trade*
		Services	*service*
2	Informality of HBs	Formal	*formal*
		Informal	*informal*
3	Accounting method	Professional accounting	*acc_prof*
		Simplified accounting	*acc_sim*
		Personal record/ notes	*acc_pers*
		No accounts	*acc_no*
4	Type of premise	Unstable premise	*premi_unstab*
		Premise at home	*premi_home*
		Professional premise	*premi_prof*
5	Whether a HB has wage earners or not	No	*wearner_no*
		Yes	*wearner_yes*
6	Total size of HB	1 person	*size=1*
		2 persons	*size=2*
		3 to 6 persons	*size3-6*
		More than 6 persons	*size>6*
7	Quintile of monthly turnover of HB	First quintile	*turnov_q1*
		Second quintile	*turnov_q2*
		Third quintile	*turnov_q3*
		Fourth quintile	*turnov_q4*
		Fifth quintile	*turnov_q5*
8	Quintile of monthly value-added of HB	First quintile	*va_q1*
		Second quintile	*va_q2*
		Third quintile	*va_q3*
		Fourth quintile	*va_q4*
		Fifth quintile	*va_q5*
9	Quintile of monthly equipment costs of HB	First quintile	*equip_q1*
		Second quintile	*equip_q2*
		Third quintile	*equip_q3*

		Fourth quintile	*equip_q4*
		Fifth quintile	*equip_q5*
10	Main customer of HB	Public or para-public sector	*cust_pub*
		Enterprises	*cust_enterp*
		Household businesses	*cust_hb*
		Individuals	*cust_ind*
11	Main supplier of HB	Public or para-public sector	*supp_pub*
		Enterprises	*supp_enterp*
		Household businesses	*supp_hb*
		Individuals	*supp_ind*
12	Age of HBs	Less than 3 years	*agehb<3*
		From 3 to 7 years	*agehb3-7*
		From 8 to 15 years	*agehb8-15*
		More than 15 years	*agehb>15*

Characteristics of HB head

13	Gender of HB heads	Male	*male*
		Female	*female*
14	Age of HB head	Less than 40 years old	*age<40*
		From 40 to 45 years old	*age40-45*
		From 46 to 55 years old	*age46-55*
		More than 55 years old	*age>55*
15	Educational level of HB head	Primary school	*edu_prim*
		Secondary school	*edu_sec*
		High school	*edu_highsc*
		University or more	*edu_univ*
16	Apprenticeship of HB head	Technical school	*tech_school*
		On-job training	*on_job_train*
		Self-trained	*self_train*
		Others	*apprent_other*
17	Immigration status of HB head	Has lived in this province/ city for at least 10 years	*lived>10y*
		Has migrated from another city to this province/ city for less than 10 years	*immig_city*
		Has migrated from a countryside to this province/ city for less than 10 years	*immig_rural*

18	Residence status of HB heads	Registered in the same province (KT1-KT2)	*KT1-2*
		Temporary registration for a period of 6 months and more (KT3)	*KT3*
		Temporary registration for a period of less than 6 months (KT4)	*KT4*
		No registration at the destination	
19	Why you do not register your HB?	Not obligatory	*ynotreg_uncompul*
		Do not know if it has to be registered	*ynotreg_0know*
		Other reasons (e.g., too complicated procedures, too expensive)	*ynotreg_other*
20	Reasons to set up or manage this HB	Did not find a salaried job in an enterprise	*a4d_nojob_ent*
		Did not find a salaried job in a HB	*a4d_nojob_hb*
		To earn higher income	*a4d_inc*
		To be independent	*a4d_indep*
		To follow family tradition	*a4d_fam_trad*
		Others	*a4d_other*
21	The main reason to choose this business activity	Family tradition	*g1_fam_trad*
		The profession that you know	*g1_prof_know*
		Have higher profit than other activities	*g1_higher_profit*
		More stable returns than other products	*g1_stabler_return*
		Others	*g1_other*
22	Willingness to register your HB	Yes	*wilg_reg_yes*
		No	*wilg_reg_no*
		Do not know	*wilg_reg_0know*
		Already registered	*registerd*
23	Do you think there is a future for a HB like yours?	Yes	*futur_yes*
		No, and think of changing activity	*nofutur_change*
		No, but do not think of changing activity	*nofutur_notchange*

Source: HB&IS HCMC (2009); author' calculations.

Table 19. Mean Tests for the Non-responses to Open-ended Questions

Variable	Non-respondents	Respondents	Difference
	Disadvantages		
Total size	1.48	1.55	-0.07
Value-added	6507.52	5458.01	1049.509
Educational attainment	7.79	7.84	-0.04
	Advantages		
Total size	1.38	1.63	-0.25**
Value-added	5291.29	6030.65	-739.36
Educational attainment	7.78	7.85	-0.08

*Note: **: significant at 5% level*

Source: HB&IS HCMC (2009); author' calculations.

Table 20. Positions of Clusters in the Dimensional Space – Advantages of Informal HBs

Clusters		Coordinates		Test-values	
N°	Size	1	2	1	2
1	320	-0.02	0.12	-1.11	5.7
2	63	0.15	0.26	1.66	3.03
3	29	-0.68	0.13	-4.93	0.99
4	16	2.49	-1.93	13.28	-10.53
5	5	-4.36	-5.71	-12.8	-17.15

Source: HB&IS HCMC (2009); author' calculations.

Table 21. Characteristic Responses of the Clusters – Cluster Analysis on Advantages of Informal HBs

Cluster N°	Characteristic responses
1	Customers (stable, many familiar clients)
2	Work load is light and simple, quite comfortable
3	Can afford family's living
4	Do not have to pay taxes
5	Helps and supports from friends, local authority and family

Source: HB&IS HCMC (2009); author' calculations.

Table 22. Categorical Characteristics of the Clusters – Cluster Analysis on Advantages of Informal HBs

Characteristic categories	v.test	Percentages			Weight	Prob.
		cla/mod	mod/cla	global		
Cluster 1				66.05	286	
a4d_to_be_independen	3.71	75.65	51.05	44.57	193	0.0001
Premise_home	2.95	73.02	54.9	49.65	215	0.0016
equip_quin4	2.8	78.49	25.52	21.48	93	0.0025
Cluster 2				19.63	85	
Premise_unstable	2.42	26.32	47.06	35.1	152	0.0077
Customer_enterprise	2.35	50	8.24	3.23	14	0.0093
Cluster 4				3.7	16	
equip_quin2	2.99	9.17	62.5	25.17	109	0.0014
a4d_to_get_better_in	2.69	11.29	43.75	14.32	62	0.0035
G1_higher_profit	2.39	16.67	25	5.54	24	0.0085

Source: HB&IS HCMC (2009); author' calculations.

References

Aydin, E., Hisarciklilar, M., & Ilkkaracan, I. (2010). Formal versus Informal Labor Market Segmentation in Turkey in the Course of Market Liberalization. *Topics in Middle Eastern and North African Economies*, *12*, 1-43.

Bacchetta, M., Ernst, E., & Bustamante, J. P. (2009). *Globalization and Informal Jobs in Developing Countries*. Geneva: International Labor Organization and World Trade Organization.

Bargain, O., & Kwenda, P. (2011). Earnings Structures, Informal Employment, and Self-Employment: New Evidence from Brazil, Mexico, and South Africa. *Review of Income and Wealth*, *57*(s1), S100-S122.

Benzécri, J. P. (1973). *L'analyse des Données: l'Analyse des Correspondances*. Paris: Dunod.

Blunch, N. H., Canagarajah, S., & Raju, D. (2001). The informal sector revisited: A synthesis across space and time. *World Bank Social Protection Discussion Papers*, *119*.

Bryman, A. (2006). Integrating Quantitative and Qualitative Research: How is it Done?. *Qualitative Research*, *6*(1), 97-113.

Cling, J. P., Nguyen, H. C., Razafindrakoto, M., & Roubaud, F. (2010a). How Deep Was the Impact of the Economic Crisis in Vietnam? A Focus on the Informal Sector in Hanoi and Ho Chi Minh City (World Bank Working Paper No. 61676). Retrieved from website: http://basepub.dauphine.fr/bitstream/handle/123456789/7476/616760WP0P10761010ENG0Final0version.pdf?sequence=1

Cling, J.-P., Nguyen, T. T. H., Nguyen, H. C., Phan, T. N. T., Razafindrakoto, M., & Roubaud, F. (2010b*). The informal sector in Vietnam. A focus on Hanoi and Ho Chi Minh City*. Hanoi: Thê Gioi Publisher.

Cling J. P., Razafindrakoto, M. & Roubaud, F. (2010c). Assessing the Potential Impact of the Global Crisis on the Labor Market and the Informal Sector in Vietnam. *Journal of Economics & Development*, *38*, 16-25.

Cling J. P., Razafindrakoto, M. & Roubaud, F. (2010d). *The Informal Economy in Vietnam*. Hanoi: International Labor Organization.

Cling, J. P., Razafindrakoto, M., & Roubaud, F. (2012). To Be or Not to Be Registered? Explanatory Factors behind Formalizing Non-Farm Household Businesses in Vietnam. *Journal of the Asia Pacific Economy*, *17*(4), 632-652.

Cling, J.P., Razafindrakoto, M. & Roubaud, F. (2014). *Segmentation and informality in Vietnam: A survey of the literature* (No. DT/2014/14).

Cunningham, W. V., & Maloney, W. F. (2001). Heterogeneity among Mexico's Microenterprises: An Application of Factor and Cluster Analysis. *Economic Development and Cultural Change*, *50*(1), 131-156.

De Soto, H. (1989). *The Other Path: The Invisible Revolution in the Third World*. New York: Harper and Row.

Dinh, D., Hoang, K., & Nguyen, V. T. (2001). Vietnamese Word Segmentation. In *Proceedings of the Sixth Natural Language Processing Pacific Rim Symposium (NLPRS2001), Tokyo (Japan)* (pp. 749-756).

Economica Vietnam. (2013). *The Non-Farm Household Business Sector in Vietnam*. Retrieved from: http://www.economica.vn/Portals/0/MauBieu/fd921a86edc2a4407dc0ad2543d0e92d.pdf.

Fields, G. S. (1990). Labor Market Modeling and the Urban Informal Sector: Theory and Evidence [Electronic version]. In D. Turnham, B. Salomé, & A. Schwarz (Eds.), *The Informal Sector Revisited* (pp. 49-69). Paris: Organization for Economic Co-operation and Development.

Guttman, L. (1954). Some necessary conditions for common factor analysis. *Psychometrika*, *19*, 149–161.

Harati, R. (2013). Heterogeneity in the Egyptian Informal Labor Market: Choice or Obligation? (CES Working Paper No 2013/32). Retrieved from Centre d'Economie de la Sorbonne website: ftp://mse.univ-paris1.fr/pub/mse/CES2013/13032.pdf.

Harris, J. R. & Todaro, M. P. (1970). Migration, unemployment, and development: A two-sector analysis. *American Economic Review*, *60*(1), 126-142.

Hughes, C. (2012). Qualitative and Quantitative Approaches. Retrieved from: http://www2.warwick.ac.uk/fac/soc/sociology/staff/hughes/researchprocess/quantitative_and_qualitative_approaches.docx.

Husson, F., Josse, J., & Pages, J. (2010). Principal Component Methods - Hierarchical Clustering -Partitional Clustering: Why Would We Need to Choose for Visualizing Data. *Applied Mathematics Department*.

Kaiser, H. F. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement, 20*, 141–151.

Kaiser, H. F. (1970). A second generation Little Jiffy. *Psychometrika, 35*, 401–417.

Porta, R. L., & Shleifer, A. (2008). *The unofficial economy and economic development* (No. w14520). National Bureau of Economic Research.

Lebart, L., & Mirkin, B. (1993). Correspondence Analysis and Classification. *Multivariate Analysis: Future Directions*, *2*, 341-357.

Lebart, L. M., Morineau, A. A., & Piron, M. (2006). *Statistique Exploratoire Multidimensionnelle* (4th ed.). Paris: Dunod.

Lebart, L., & Piron, M. (2012). Exploring Numerical and Textual Data in Practice with Dtm-Vic. Retrieved from: http://www.dtmvic.com/doc/Manual_DtmVic_format_A5.pdf.

Lebart, L., Salem, A., & Berry, L. (1998). *Exploring Textual Data*. Dordrecht: Kluwer.

Lewis, W. A. (1954). Economic development with unlimited supplies of labor. *Manchester School*, *28*(2), 139-191.

Maloney, W. F. (2004). Informality revisited. *World Development*, *32*(7), 1159-1178.

Nguyen, C. T., Nguyen, T. K., Phan, X. H., Nguyen, L. M., & Ha, Q. T. (2006). Vietnamese Word Segmentation with CRFs and SVMs: An Investigation. In *Proceedings of the 20th Pacific Asia Conference on Language, Information and Computation (PACLIC 2006)*. Paris: Organization for Economic Co-operation and Development.

Nguyen, T. V., Tran, H. K., Nguyen, T. T., & Nguyen, H. (2006). Word Segmentation for Vietnamese Text Categorization: An Online Corpus Approach. In *the 4rd IEEE International Conference in Computer Science*.

Oostendorp, R. H., Tran, Q. T. & Nguyen, T. T. (2009). The Changing Role of Non-Farm Household Enterprises in Vietnam. *World Development*, *37*(3), 632-644.

Piau, C. (2004). *Quelques Expériences sur la Formulation des Questions d'Enquête: à Partir du Matériau « Aspirations et Conditions de Vie des Français*. » CRÉDOC.

Portes, A., Castells, M. & Benton, L. A. (Eds.) (1989). *The Informal Economy: Studies in Advanced and Less Developed Countries*. Baltimore MD: The John Hopkins University Press.

Rakotomanana, F. H., & Rakotoasimbola, J. J. (2015). Les Cause de l'Informalité à Madagascar : Avantages et Inconvénients de Travailler dans le Secteur Informel selon les Travailleurs Indépendents – Application de la Méthode d'Analyse des Données Textuelles.

Rand, J., & Torm, N. (2012). The benefits of formalization: Evidence from Vietnamese manufacturing SMEs. *World Development*, *40*(5), 983-998.

Razafindrakoto, M. & Nguyen, H.C. (2010). Household businesses and the informal sector in Hanoi and Ho Chi Minh City: first results from a qualitative survey. GSO/IRD Working Paper. Hanoi: GSO/IRD.

Razafindrakoto, M., Roubaud, F., & Wachsberger, J. M. (2012). Travailler dans le Secteur Informel : Choix ou Contrainte? Une Analyse de la Satisfaction dans l'Emploi au Vietnam. In *L'économie Informelle dans les Pays en Développement*, Cling, J. P., Lagrée, S., Razafindrakoto, M., Roubaud, F., eds., Edition AFD, collection conférence et séminaire, 47-66.

Schuman, H., & Presser, S. (1996). *Questions and Answers in Attitude Surveys: Experiments on Question Form, Wording, and Context*. SAGE Publications, Incorporated.

Taussig, M., & Pham, T. T. H. (2004). *Private Enterprise Formality and the Role of Local Governance.* ADB Making Markets Work Better for the Poor, Discussion Paper No. 02. Hanoi: Asian Development Bank.

Teddlie, C., & Tashakkori, A. (2008). *Foundations of Mixed Methods Research: Integrating Quantitative and Qualitative Approaches in the Social and Behavioral Sciences*. SAGE Publications, Incorporated.

Tran, Q. T., Pham, T. T., Ngo, Q. H., Dinh, D., & Collier, N. (2007). Named Entity Recognition in Vietnamese Documents. *Progress in Informatics Journal, 5,* 14-17.

Vu, H. N., Le, T. T. L., & Nguyen, T. N. A. (2010). Factors Impinging on the Development of Craft Village Enterprises: the Case of Van Phuc Silk Village. *Economic Development Review (HCMC University of Economics), 191,* 9-13.

Wanner, L. (2004). Introduction to Clustering Techniques. *International Union of Local Authorities.* Retrieved from: https://www.iula.upf.edu/materials/040701wanner.pdf.

Word. (2013). In *Oxford Dictionary*. Retrieved May 26, 2013, from http://oxforddictionaries.com.

Yelland, P. M. (2010). An Introduction to Correspondence Analysis. *The Mathematica Journal, 12.*

Chapter 2

Returns to Higher Education in Vietnam: Who Benefit the Most?

Abstract

Higher education has been playing a key role in the development strategy of Vietnam's government. Numerous studies have investigated the return to education in Vietnam and its evolution over the transition period. Yet, this is the first research paper focusing on the heterogeneity of returns to higher education across Vietnamese population. Based on Vietnam Labor Force Survey 2007, results from OLS, IV, Heckman, Quantile Regression and Heterogeneous Treatment Effect models all reveal relatively large returns that greatly vary across individuals. Evidence for Negative Selection Hypothesis is found: those who are the least likely to obtain higher education tend to benefit the most from it.

Acknowledgements

The current version of this paper has benefited from valuable comments at the 8th Vietnam Economist Annual Meeting (VEAM), Thai Nguyen, Vietnam; at the 2nd Sustainability Workshop in Rouen, France; and at the DIAL-CEPED Workshop on "Inequality and Vulnerability in Vietnam" in Paris, France.

1. Introduction

The transition period in Vietnam has started in 1986 with a sequence of economic reforms (i.e., *Doi Moi,* or *Renovation*), including agricultural de-collectivization, land reforms, increasing autonomy of state-owned enterprises, enhancement of private sector and trade liberalization (World Bank, 1998). Throughout the process, the demand for skills has been augmenting, due to inter-industry employment changes, capital accumulation and skill-biased technical change (World Bank, 2008). Consistently, various studies have investigated the return to education (or *skill premium*) in Vietnam and its evolution in the transition process, showing that it has followed an upward trend over the last decades, reaching a relatively high level recently (Gallup, 2002; Arcand et al., 2005; Liu, 2006; Doan & Gibson, 2010; Doan, 2011; Oostendorp & Doan, 2013; Phan & Coxhead, 2013; Nguyen, 2014 among others).[36] Vietnam's Socio-Economic Development Strategy for the period of 2011-2020 considers promoting human resources/skills development (particularly skills for modern industry and innovation) the first of three 'breakthrough areas.'[37]

Nonetheless, Vietnam's higher education system has not been fully equipped to meet up with such increasing need for highly qualified labor.[38] Despite a considerable budget share dedicated to higher education,[39] both access (i.e., equality) and quality concerns remain. On the one hand, although the current boom of colleges and universities in Vietnam has, to some extent, broadened the coverage of higher education, its enrolment rate is still low (Hayden & Lam, 2015). Facing a lack of available places, young people from rural areas and poor backgrounds have less chance to be included among enrolments. This means a loss of opportunities for the country, especially when some of the most brilliant cannot enter the university gate (World Bank, 2008). On the other hand, higher education institutions' quality has been questioned. According to Vietnam's Former Minister of Education and Training Pham Vu Luan, "Vietnam is now suffering from an excess of low quality universities and a lack of high quality ones."[40] The quality problems bring out academic failure, lack of interest in staying school and lower future earnings (Di Gropello, 2006). This implies that schooling quality and returns to education are closely related terms.

It is at times believed that educational access and educational quality are trade-offs. Di Gropello (2006) argues that at tertiary level, this scenario occurs frequently as public investment focuses on excellent universities that are accessible only to the elites. By contrast, case studies show

[36] Return to education is understood as the (relative) change in income or wage earned resulting from having an additional amount of education.

[37] Source: http://dsi.mpi.gov.vn/vietnam2035/en/3/49.html

[38] Higher education is defined as education at universities or similar educational establishments, especially to degree level (Source: oxforddictionaries.com).

[39] Vietnam has spent nearly one percent of their GDP on higher education, compared to 0.15 percent, 0.53 percent and 0.32 percent for Myanmar, Indonesia and the Philippines, respectively (UNESCO, 2014).

[40] This is his answer at the assembly on 24 November 2011, in response to numerous questions mostly raising concerns that the deteriorating quality of higher education was becoming uncontrollable, especially in private institutions. Source: http://www.universityworldnews.com/article.php?story=2011120222340338

that the dual role can be accomplished in education systems that combine high-selective institutions and more open admission to the others (ibid).

The debate of quality-access seems relevant to the efficiency-equality dilemma facing Vietnam's education policy makers. Efficiency, or cost-effectiveness, represents the relationship between the education's outputs with the invested inputs. On the one hand, limited resources may require the investment in tertiary education to prioritize certain groups of individuals with the highest returns, who will then contribute much to the national budget through income taxes. Standard cost-benefit theory suggests that those who are most likely to select themselves in higher education (e.g., students from wealthy families) are the ones with the highest returns. On the other hand, equality objective induces the State to pay attention to socially disadvantaged individuals, whose capacity to reap the benefits of higher education are not surely high. If these are the case, Vietnam's government might face a tradeoff between efficiency and equality. Although the existing literature has acknowledged the heterogeneity of the skill premium in Vietnam, this subject has never been thoroughly investigated as the focus of the analysis.

This is the first study that focuses on examining the heterogeneity of the returns to higher education on individual earnings in Vietnam. As expected, the author finds relatively large returns that greatly vary across individuals. The most important finding of this paper is that, contradicting to the above concern about efficiency-equity tradeoff, those who are least likely to obtain tertiary education actually tend to benefit the most from it. This result is consistent across different models. The paper is structured as follows. Section 2 reviews the literature on returns to education in general, with a special focus on studies in Vietnam. Section 3 presents data and descriptive statistics, followed by methodologies in section 4. Section 5 and 6 provide econometric findings. Robustness checks are implemented in section 7 and the final section is devoted to discussion and conclusion.

2. Literature review

Return to higher education has been investigated in a vast literature all over the world. In general, it is found positive and significant, varying across the population (Card, 1999; Psacharopoulos & Patrinos, 2004). The world average private return to higher education is recorded at 19 percent, with the highest return in Sub-Saharan Africa (27.8 percent) and the lowest in OECD countries (11.6 percent), according to a survey by Psacharopoulos and Patrinos (2004). Astonishingly, it is documented that the estimation method makes little difference on the returns to education. Card (1999), after surveying a vast literature, concludes that the average return to education is not much below the estimate obtained from a standard human capital earnings function fit by Ordinary Least Squares (OLS).

Furthermore, numerous studies have found evidence for heterogeneity in returns to education, such as Card and Krueger (1992), Blundell et al. (2004), Heckman and Li (2004). First, it is possible that the impact of education varies according to some individual characteristics or with probability of obtaining education. Second, the returns to education might also depend on

the relative position of individuals in the earnings distribution. Regarding the first possibility, there are two opposite views. On the one hand, rational-behavioral models consider that college decision is made based on a cost-benefit consideration (Mincer, 1974), suggesting that the most 'college worthy' individuals (i.e., having the highest returns to higher education), are the most likely to select into college (Carneiro et al., 2001; Heckman & Li, 2004). On the other hand, sociological literature acknowledges numerous non-economic factors that influence higher education decision. College attainment is driven not only by rational choice but also by cultural and social norms and circumstances (Coleman, 1988). It is argued that high social background individuals might attend college without a rational economic cost-benefit analysis. Meanwhile, low social background ones must overcome major odds to complete higher education.

In fact, evidence shows that individuals with relatively disadvantageous social backgrounds, would benefit the most from going to college (Golthorpe & Jackson, 2008). Based on an innovative hierarchical linear model applied to the United States data, Brand and Xie (2010) uncover a surprising *negative selection* phenomenon, i.e., those who are the least likely to obtain a college education seem to benefit the most from it. The authors argue that these people, whose position is marked by substantial disadvantage, would have limited human, cultural and social capital, thus limited labor market prospects unless they obtain higher education. Meanwhile, the advantaged individuals could still profit from their superior resources and capacities in the absence of such a degree. Indeed, evidence suggests that programs raising the educational level of children from poor family backgrounds tend to have higher marginal returns (Kane & Rouse, 1995). Likewise, studies that use instrumental variables affecting individuals on the margin of school continuation obtain larger estimated returns than do OLS estimates.

The returns to education might also depend on the relative position of individuals in the earnings distribution. A survey of literature by Di Gropello (2006) shows that in most of the countries with available empirical evidence, returns to education increase with earnings. This is consistent with the concept of complementarity of skills, whereby the productivity of skills obtained through education is enhanced by other cognitive and non-cognitive abilities. However, this phenomenon happens only in sufficiently competitive labor markets such as Europe, the United States and whites in South Africa. Actually, for low-income and lower-middle-income countries in East Asia and the Pacific, the opposite situation occurs.

Indeed, using Instrument Variable Quantile Regression method on Chinese data, Wang (2013) finds that the returns to education decrease with earnings. The author argues that the earnings percentile could be considered as the level of individual ability,[41] thus this result implies abler individuals receive lower marginal returns to schooling than do less able individuals. Several interpretations are put forward. First, education and abilities are substitutes, suggesting the negative relationship between education and ability. Second, due to an over-emphasis on exam-taking techniques of Chinese schools, workers often have fewer skills than required, making education matter less among high-ability ones. The third explanation relates to China's restrictive mobility system (*hukou*, i.e., household registration). Less able individuals might have fewer means to move,

[41] Wang (2013) cited prior literature with the same argument: Chernozhukov et al., 2007; Arias et al., 2001; Patrinos et al., 2006.

and education could allow them to migrate to better places for better-paid jobs. Moreover, it is possible for college students to transfer their household registrations to collective household registration of the universities where they enrolled, expanding the impact of education in the lower tiers of the earnings distribution. Overall, the observed distribution pattern of returns to schooling, according to the author, implies that education could reduce earnings inequality while increasing individual earnings. In other words, education-enhancing policies could achieve efficiency and equity simultaneously.

In Vietnam, returns to education have been estimated in a number of studies. Main addressed aspects include rate of return and its evolution during the transition period, policies leading to these changes, income disparity between public and private sectors, and decomposition of gender wage gap (Nguyen, 2014). Despite the difference in magnitude, the impact of education appears to be always positive and significant, and the rates of return tend to increase during the transition period (Stroup & Hargrove, 1969; Gallup, 2002; Moock et al., 2003; Arcand et al., 2005; Doan & Gibson, 2010; Liu, 2006; Doan, 2011; Oostendorp & Doan, 2013; Phan & Coxhead, 2013; Nguyen, 2014; among others).[42]

Moock et al. (2003) serves as a benchmark study for further research on the return to education in Vietnam. Using Mincer earnings functions on the Vietnam Living Standards Survey (VLSS) 1992-1993, the authors find an average private return to another year of schooling of five percent, and to higher education of 11 percent for wage workers.[43] The study also points out a remarkable heterogeneity in the returns to education with respect to sex, level of education and institutional sector. Females experience a much higher rate of return than males. Workers in the public sector have higher private rates of returns to education than do private sector counterparts. The authors suggest that there are factors other than education, for example, Communist Party membership, that distort public sector pay.

Arcand et al. (2005) focus on the choice of instrument variables (IVs) in estimating the returns to education in Vietnam for a sample of 324 wage-earning males over the period 1992-1998, based on VLSS 1993 and 1998. The authors acknowledge that heterogeneous returns to education across the population cause different estimated results corresponding to different instruments used. Demand-side valid instruments which are uncorrelated with individual earnings are difficult to find, and supply-side instruments often encounter the problem of 'weak instruments.' Carrying out several sets of tests, the paper examines the validity, relevance and performance of different types of instruments. Obtained results suggest that many conventional IVs such as school proximity and smoking habits are indeed invalid, and that only parental education and Hausman-Taylor matrix of instruments satisfy two usual validity requirements.[44]

[42] A literature summary of the returns to education in Vietnam is provided in Table 16 (Appendix).

[43] Mincer equation: the earnings function developed by Mincer (1974), in which individual earnings is explained by their schooling, experience and its square.

[44] The HT estimator, developed by Hausman and Taylor (1981) for use with panel data, solves the problem of correlation between explanatory variables and latent individual effects. In this method, time-varying variables are used to estimate their own coefficients and to serve as IVs for endogenous time-invariant variables (ibid). It provides

One additional year of education is expected to raise hourly wage by five to seven percent according to IV estimations, which is much larger than OLS estimate at 2.6 percent.

Using VLSS data in 1992-93 and 1997-98, Liu (2006) studies the evolution in earnings of male and female wage workers with different educational levels in Vietnam. The author applies OLS Mincer earnings equations without selection correction for the 1998 round and with selection correction for the 1993 round. Results show that the return to an additional year of education for women increased slightly over the period (4.2 percent in 1993; 4.8 percent in 1998), while an opposite trend happened to men (six percent in 1993; 3.5 percent in 1998). This is possibly because more women became wage workers in the FDI[45] sector where education is more rewarded. Particularly, applying Katz-Murphy (1992)'s demand-supply method, this paper finds that earnings of workers with tertiary education markedly increased relative to workers with below-primary education, as relative demand shifted in favor of higher-educated workers over the period.

Based on the 1998, 2002, 2004 and 2008 rounds of VLSS and VHLSS[46] data, Doan and Gibsons (2010) discover a rapid rise in the return to education in Vietnam during the later economic reform, but its speed has slowed down when it reaches the global average rate (from 3.8 percent in 1998 to ten percent in 2008). Such a rising trend seems robust to the self-selection into the wage employment. The authors link this trend with "further market opening and integration into the global economy, deeper reforms, and a consequent investment boom with accelerated structural change that has generated many technical-skilled jobs in Vietnam." They also predict that the rate of returns to schooling is likely to keep rising until Vietnam becomes a universally-recognized market economy.

Doan (2011) employs a variety of estimation methods, including OLS, IV and Treatment Effect model to estimate the returns to university education with VHLSS 2008 data. The author finds that tertiary education is expected to raise individual wage rate by 68 percent, *ceteris paribus*, which is much larger than the average level of Asia (18.2 percent, Psacharopoulos & Patrinos, 2004) as well as than Vietnam's estimates derived from previous data. Doan (2011) argues that it could be related to recent labor market reforms that reward higher skilled workers in Vietnam's economy.

A recent study by Osstendorp and Doan (2013) investigates the trend in returns to education during transition, pointing out a fact that trade liberalization has not only wage effect but also employment effect. The authors argue that preceding studies that ignore the latter effect fail to obtain the total/unconditional returns to education. They first adopt a modified Mincer equation, which allows for nonlinear returns to education by using a third-order polynomial of the education variable, and corrects for bias induced by self-selection into different job categories. It

consistent and efficient estimates of the coefficients associated with time-invariant variables despite the absence of external instrument (Arcand et al., 2004).

[45] Foreign Direct Investment.

[46] Vietnam Household Living Standards Surveys take into account the advantages and limits of VLSS 1992-93 and 1997-98.

is found that the Mincer returns to education for wage workers underwent an increase over the period 1998-2006. Next, a multinomial logit model for sector choice is applied, showing negative employment effect. Wages are found to be higher in industries with relatively more low-skilled workers. The authors then re-estimate the Mincer equation for each sector separately, based on which unconditional returns to education are estimated, comprised of two components: wage effect and employment effect. They conclude that the presence of negative employment effects implies that the Mincer returns actually overestimate the unconditional returns.

Such an increase in the returns to education over the transition period (1993-2008) is also confirmed in Phan and Coxhead (2013) with VHLSS data. The paper emphasizes that return to college education increased the most. Heckman method is applied to correct for workers' self-selection into different employment groups, especially the State sector. Although OLS estimates are found biased upward compared to Heckman's, the difference is inconsiderable, and the trends are the same.

Nguyen (2014) is probably the first study on returns to education using Labor Force Survey (LFS 2012) data. Applying random effects estimation for Mincer equations and a Heckman correction model, the author shows that the returns to education are generally low, especially for tertiary level. Holding a university or college degree could raise individual earnings by 68 percent, compared to those with no education. The study also shows different levels of returns to education by gender, formality, sector and types of ownership. In general, women have higher rates of return than men across all education levels. Particularly in the informal sector, the expected return to higher education is 29.4 percent for males and 63.5 percent for females. Public sector appears to value tertiary qualification the most, at 96 percent. Agriculture sector has the lowest return to higher education, at 43.9 percent, compared to over 80 percent in Industry and Services. Banking, finance, insurance; and science, technology and professional activities are the two subsectors that reward higher education the most, at 224 percent and 190 percent, respectively.

In summary, existing studies on returns to education, including higher education, in Vietnam have confirmed the presence of wage/earnings premium associated with education, which follows an upward trend for the recent decades. Such an increasing trend could be considered as the consequence of technology- or skill- biased transition/trade liberalization process, in contrast to the earlier experience of the East Asian Newly Industrialized Countries during the 1960s and 1970s (Oostendorp & Doan, 2013). These studies have more or less acknowledged the heterogeneity in returns to education via separate analyses on different groups. Nonetheless, they all estimate the rates of return at mean level. This research work would make an important contribution to the literature, by investigating the variation of returns to higher education in Vietnam across the population, based on their propensity of obtaining education or their relative position in the earnings distribution.

3. Data and descriptive statistics

3.1. Data

The analyses in this study are based on the LFS 2007, a nationally representative dataset. Individuals of at least 25 years old, having completed schooling and with available earnings information are included in the analysis.[47] The hourly earnings are computed based on monthly earnings and actual number of working hours per week.[48] In the LFS 2007, the question on monthly earnings captures all salary/wage, bonuses and subsidies, so there is no possibility to single out individual wage rate. Only the main job in the last seven days is counted. *Experience* variable is approximated by individual age.

The higher education variable is defined using three questions in the LFS 2007 questionnaire. Question 9 asks if the interviewee is attending, or has attended or never attended school. Evidently, the earnings analysis in this paper should cover only those who have completed their schooling. Question 10 provides information on the highest grade of general school ('lop pho thong') completed by the interviewee. This variable is only available in LFS 2007. Question 11 shows the highest technical/ professional qualification that one has obtained. Since the survey does not provide the exact number of years of schoolings, higher education can only be presented by a dummy variable. To isolate the impact of tertiary education from the cumulative effects of lower education levels, only those who have completed high school (i.e., twelve years of general education) are kept for the analyses. In short, the focus of this research is the earnings gap between high school leavers and university (and above) graduates.

LFS 2007 is chosen because of at least three reasons. First, it has the largest sample size (173,000 households) among all the VLSS, VHLSS and LFS that have been conducted so far.[49] Large sample size, on the one hand, offers more precise estimates. On the other hand, that facilitates the disaggregation into sub-samples. Second, as mentioned in the literature review, almost all of previous studies have exploited the VLSS or VHLSS data. To the best of my knowledge, this is the first paper that examines the returns to schooling based on LFS 2007.

Third, the education variable is one important advantage of the LFS 2007 compared to other LFS rounds (**Table 17**, Appendix). LFS 2007 (and LFS 2012) enables us to distinguish those who have finished their schooling from those who are currently attending school. Moreover, LFS 2007 is the only LFS survey that allows to identify those who have completed high school. LFS 2009, for example, provides no information on the educational level of individuals. Although LFS 2010, 2011 and 2012 contain such information, there exists a blurred distinction of those who have

[47] By the age of 25, normally in Vietnam one should have finished her university education. Social sciences bachelor programs often last four years, engineering programs five years and medical courses the longest – six years.

[48] Hourly earnings is approximated by following this formula: hourly earnings = (monthly earnings*7)/(weekly working hour*30).

[49] The sample sizes (number of households) for some other surveys are: 4,800 (VLSS 1992-93); 6,000 (VLSS 1997-98); 75,000 (VHLSS 2002); 46,000 (VHLSS 2004, 2006 and 2008); 18,000 (LFS 2009); 123,000 (2010, second round); 155,014 (LFS 2011) and 102,651 (LFS 2012). Source: General Statistics Office.

finished twelve years of general education and those who have not. In Vietnam, those who have obtained a short-term training or graduated from trade vocational school might have just completed nine years of general education.[50]

One may argue that one advantage of VLSS over LFS is that VLSS supports panel data analyses, while LFS does not. Veritably in this particular setting, higher education as a treatment status offers little variation from year to year. Neither LFS nor VHLSS provide exact number of years of higher education. Moreover, panel data analyses could introduce attrition bias and sample selection bias (Oostendorp & Doan, 2013). Given limited time span, those who change their treatment status are either newly graduates or officials taking in-service training ('tai chuc') of usually low quality. The returns to higher education for these groups might be specific. Furthermore, measurement error in schooling tends to be higher in fixed-effect model than in cross-sectional estimator (Ashenfelter & Zimmerman, 1997; Belzil, 2007). These are the reasons why cross-sectional analysis is chosen over panel one in this research work.

3.2. Descriptive Statistics

Table 1 provides a global picture of higher education coverage in the whole sample of those who have completed their schooling. Overall, higher education graduates made up 9.4 percent of the population, and this proportion substantially varied across different socio-economic groups. In line with conventional ideas, the ethnic minority, females, rural residents, agricultural workers were the least likely to complete a university education. Moreover, there seemed to be a gradual progress of tertiary education coverage from generation to generation. While the oldest (at least 56 years old) rarely obtained a college qualification (6.7 percent), the younger generations tended to go to university more often (10 percent for 41-45 years old and 11.6 percent for 25-40 years old). Among all geographical regions, the rate of higher education graduates appeared to be the lowest in Mekong River Delta (5.4 percent), while the highest in Red River Delta (12.2 percent). Services workers were outstandingly the best educated among four sectors, with up to 28.5 percent holding a college degree, compared to only 0.7 percent in agriculture, 6.7 percent in manufacturing and 4.8 percent in trade. Finally, public sector workers were undoubtedly the most qualified (43.8 percent), far exceeding all other sectors, even workers in foreign enterprises (13 percent).

We then examine the situation of higher education in our restricted sample (**Table 2**). Overall, about one third of high school graduates in our sample went on for higher education. This figure is over three times higher than that in the above statistics, owing to the fact that our restricted sample includes only 25-years-old-and-over individuals who have completed twelve years of general schooling with available earnings information. This rate seems more or less stable among various age groups, genders and regions, while significantly fluctuates across ethnic groups, urban/rural areas, industries and institutional sectors. To some extent, the between-group variation shrinks, probably because this restricted sample is less heterogeneous than the non-restricted sample.

[50] References: http://tuyensinhtoanquoc.com/thong-tin-tuyen-sinh-2/tot-nghiep-lop-9-duoc-hoc-tiep-len-trung-cap-nghe-chinh-quy/; http://tuyensinhtrungcap.vn/hoc-het-cap-2-co-hoc-trung-cap-duoc-khong/

Table 1. Higher Education Coverage in Vietnam (unrestricted sample)

	Proportion having completed higher education		
	Mean (%)	**Std. Err.**	**Obs.**
Whole sample	**9,4**	**0,000**	**420066**
Age			
25-40	11,6	0,001	355566
41-55	10,0	0,001	
>=56	6,7	0,001	
Ethnicity			
Kinh	9,9	0,000	420058
Minor	5,1	0,001	
Gender			
Male	10,2	0,001	420066
Female	8,6	0,001	
Area			
Urban	13,1	0,001	420066
Rural	2,6	0,000	
Region			
Red River Delta	12,2	0,001	420066
Northern Midlands and Mountain Areas	10,5	0,001	
Northern Central and Central Coastal Areas	10,3	0,001	
Central Highlands	8,9	0,002	
South East	8,9	0,001	
Mekong River Delta	5,4	0,001	
Industry			
Agriculture	0,7	0,000	420066
Manufacturing	6,7	0,001	
Trade	4,8	0,001	
Services	28,5	0,001	
Institutional sector			
Public sector	43,8	0,002	420066
Foreign enterprise	13,0	0,004	
Domestic enterprise	14,7	0,002	
Formal household business	3,9	0,001	
Informal household business	1,5	0,000	
Agriculture	0,7	0,000	
Others	6,4	0,001	

Source: LFS 2007 (GSO), author's calculation.

Table 2. Higher Education Coverage in Vietnam (restricted sample)

	Proportion having completed higher education		
	Mean	**Std. Err.**	**Obs.**
Whole sample	**0.358**	**0.002**	**85592**
Age			
25-40	0.362	0.002	85592
41-55	0.351	0.003	
>=56	0.373	0.007	
Ethnicity			
Kinh	0.360	0.002	85589
Minor	0.331	0.006	
Gender			
Male	0.355	0.002	85592
Female	0.362	0.002	
Area			
Urban	0.383	0.002	85592
Rural	0.222	0.004	
Region			
Red River Delta	0.350	0.003	85592
Northern Midlands and Mountain Areas	0.340	0.003	
Northern Central and Central Coastal Areas	0.374	0.003	
Central Highlands	0.358	0.006	
South East	0.366	0.005	
Mekong River Delta	0.370	0.005	
Industry			
Agriculture	0.091	0.003	85539
Manufacturing	0.227	0.003	
Trade	0.159	0.003	
Services	0.518	0.002	
Institutional sector			
Public sector	0.571	0.002	85592
Foreign enterprise	0.325	0.011	
Domestic enterprise	0.334	0.005	
Formal household business	0.119	0.003	
Informal household business	0.076	0.002	
Agriculture	0.091	0.003	
Others	0.359	0.019	

Source: LFS 2007 (GSO), author's calculation.

Not surprisingly, Kinh people were more inclined to go to college than ethnic minority (36 percent and 33.1 percent, respectively). Also, the coverage of higher education in the cities was nearly twice as much as that in the countryside (38.3 percent versus 22.2 percent). On average, services workers were the most educated (51.8 percent), followed by manufacturing (22.7 percent) and trade (15.9 percent). Agricultural workers, as expected, were the least likely to obtain a higher

education diploma (9.1 percent). In line with conventional perception, public sector workers had the highest average probability of having completed higher education (57.1 percent) among all institutional sectors. Workers in domestic and foreign enterprises were approximately at the sample average level. Meanwhile, household business sector was marked with an extremely less educated labor force, with the informal units having the least educated workers (7.6 percent).

Table 3 provides a comparison of those who obtain higher education and those who do not in terms of demographic as well as current employment characteristics. In general, the former group had a more privileged background and enjoy better working conditions than the latter.

Table 3. Descriptive Comparison between Non-Treated and Treated Groups

Variable	Non-treated			Treated			Difference
	Mean	Std. Err.	Obs.	Mean	Std. Err.	Obs.	
Age	39.924	0.040	54939	39.338	0.057	30653	0.586***
Ethnic minority	0.075	0.001	54938	0.067	0.001	30651	0.008***
Female	0.467	0.002	54939	0.475	0.003	30653	-0.008**
Rural	0.190	0.002	54939	0.097	0.002	30653	0.093***
Father's income	7.060	0.166	2778	9.987	0.242	2000	-2.926***
Mother's income	6.218	0.351	1500	8.744	0.447	1014	-2.526***
Father completing higher education	0.079	0.003	6222	0.270	0.007	4218	-0.190***
Mother completing higher education	0.024	0.002	5753	0.126	0.006	3473	-0.102***
Wage workers	0.508	0.002	54905	0.901	0.002	30628	-0.393***
Informal employment[51]	0.495	0.002	54939	0.099	0.002	30653	0.396***
Public sector	0.332	0.002	54939	0.793	0.002	30653	-0.461***
Hourly earnings	7.699	0.039	54939	11.763	0.049	30653	-4.064***

Source: LFS 2007 (GSO), author's calculation.

Demographic characteristics: The two groups were tightly comparable in terms of age and gender structure. The average age of both groups was about 39-40 years old. Females constituted around 47-48 percent of both samples. However, those who obtain a tertiary education were a little less likely to belong to ethnic minority than those without (6.7 percent versus 7.5 percent, respectively). The non-treated group were twice as probable as the treated group to live in rural areas (19 percent and 9.7 percent, respectively). The latter's family background was also more favorable, with wealthier and much more educated parents. Particularly, while up to 27 percent of fathers and 12.6 percent of mothers of college graduates had also completed a tertiary education, merely 7.9 percent of fathers and 2.4 percent of mothers of their non-treated counterparts held a higher education diploma. This probably indicates an inter-generational human capital transmission in Vietnam as mentioned in Arcand et al. (2004).

Job characteristics: Unsurprisingly, the better educated group also tended to acquire more positive employment characteristics. First, almost all of them were wage workers (90.1

[51] Informal employment includes all jobs in the informal sector and informal jobs in the formal sector. The informal sector is defined as all private unincorporated enterprises that produce at least some of their goods and services for sakes or barter and are not registered, excluding all farming activities (Cling et al., 2010). The informal jobs in the formal sector are defined in this chapter as those without social security coverage.

percent), compared to only a half for the control group. While a half of high school leavers were engaging in informal jobs, only 10 percent of college graduates found themselves in this vulnerable employment category. A majority of college graduates worked in the public sector (79.3 percent), whilst only one third of high school leavers did so. In Vietnam, public sector employment is highly appreciated for their stability and social status according to social norms. The average hourly earnings of the treated group were substantially larger than those of the non-treated (11.76 compared to 7.69 per hour, in thousand Vietnam dongs).

A further investigation into the hourly earnings is presented in **Table 4**. An average worker in our sample earned about nine thousand Vietnam dongs for each working hour. There was a wide range of hourly income across different sub-samples. On average, prime-aged workers appeared to earn the most (10.2) among three age groups. Kinh people had higher earnings than other ethnic groups (9.2 versus 8.5). Males were better paid than females (9.6 versus 8.6). In general, rural workers were less remunerated than their urban counterparts (7.0 versus 9.6), and workers living in Northern Central and Central Coastal Areas, Northern Midlands and Mountain Areas and Red River Delta had the lowest income among all regions. In line with conventional wisdom, informal sector and agriculture were the two sectors where workers earn the least (6.7 and 5.7, respectively). Results from **Table 3** and **Table 4** suggest that the likelihood of obtaining higher education had some correlation with earnings. This statement was particularly relevant to industry and institutional sector groups.

Table 4. Average Hourly Income in Different Groups

		Mean hourly income (in thousand dongs)		
		Mean	**Std. Err.**	**Obs.**
Whole sample		**9.153**	**0.031**	**85625**
Age				
	25-40	8.405	0.035	
	41-55	10.209	0.057	85625
	>=56	8.777	0.176	
Ethnicity				
	Kinh	9.207	0.033	85622
	Minor	8.469	0.082	
Gender				
	Male	9.641	0.046	85625
	Female	8.603	0.041	
Area				
	Urban	9.552	0.035	85625
	Rural	7.004	0.057	
Region				
	Red River Delta	8.663	0.063	
	Northern Midlands and Mountain Areas	8.619	0.048	
	Northern Central and Central Coastal Areas	8.334	0.045	85625
	Central Highlands	10.562	0.162	
	South East	11.554	0.143	

	Mekong River Delta	9.760	0.099	
Industry				
	Agriculture	5.652	0.113	
	Manufacturing	8.815	0.062	85572
	Trade	8.339	0.097	
	Services	10.175	0.039	
Institutional sector				
	Public sector	10.486	0.030	
	Foreign enterprise	10.743	0.218	
	Domestic enterprise	10.293	0.168	
	Formal household business	9.233	0.141	85625
	Informal household business	6.679	0.064	
	Agriculture	5.652	0.113	
	Others	8.982	0.246	

Source: LFS 2007 (GSO), author's calculation.

Overall, these descriptive results imply that individuals from lower social background were less likely to attend college. Moreover, they might suggest that an OLS regression which ignores individual characteristics and family background may produce a biased estimate of the returns to higher education in Vietnam. Refined assessment of an empirical strategy in estimating the earnings premium associated with tertiary qualification is presented in the next section.

4. Methodology

4.1. Potential bias of OLS estimation on returns to education

In a non-experimental context, the main challenge to the causal effect identification is that one cannot observe the earnings of the same individual in both states: with and without going to college. A naive bivariate OLS regression of individual income on her schooling might suffer from several sources of bias as follows:

(1) Omitted variable bias (Pre-treatment heterogeneity bias). There are some factors that drive both educational decision of individuals and their earnings in the labor market such as gender, ethnicity, region and social/family background. For example, people with superior family resources and status, who would earn more in the labor market *ceteris paribus*, are also more likely to obtain more education. The omission of family background thus possibly leads to an overestimation of the returns to schooling.[52]

(2) Measure errors. While measurement error in earnings only worsens precision of the model, measurement error in schooling causes attenuation bias. In fact, prior research has generally reported the reliability of self-declared schooling at only 90 percent (Card, 1999).

[52] Reverse causality is unlikely a threat to causal inference in this case because only those who have completed education are included. As their earnings are measured at the current moment, it is not affected by the education decision in the past.

Fortunately, such a downward bias could be balanced with aforementioned ability upward bias.

(3) **General equilibrium effect (Violation of Stable Unit Treatment Value Assumption - SUTVA)**. There is a possibility that an excessive number of highly qualified workers in the economy would reduce monetary reward from education in general, and induce people to study more not to lag behind. In fact, Vietnam has recently witnessed a boom of universities and tertiary faculties without restrictive quality controls, which is probably responsible for unemployment and low earnings among the graduates. Therefore, it should be kept in mind that this research is restrained within private impact of education.

(4) **Heterogeneous treatment effects**. Simple OLS imposes homogeneous returns. It is in essence a weighted average of heterogeneous effects, some of which should be higher, while others should be lower, than the population average (Angrist & Krueger, 1999).

If the impact of education varies according to some individual characteristics or with treatment probability, conventional estimates of returns to schooling might be biased. Imbens and Angrist (1994) state that an IV estimate only recovers a local treatment effect, i.e., the effect on those whose educational decision is affected by the instrument ('compliers'). Moreover, Heckman and Li (2004) show that OLS gives a downward biased estimate and IV produces an upward biased estimate of the average treatment effect (ATE).[53] The disparity between IV and OLS estimates depends on the extent that instruments affect schooling decision at different education levels due to heterogeneous returns to schooling (Card, 1999). For example, IV estimates based on instruments which influence schooling decisions of children from relatively disadvantageous family background (e.g., lower parental education, income, assets) tend to be higher than the OLS estimates.

4.2. Empirical Strategies

4.2.1. Ordinary Least Square

We first start with homogeneous treatment effect models. As a benchmark estimation, the private returns to education are first estimated in the conventional Mincer earnings function (Mincer, 1974) as follows:

$$Y = \alpha + \beta EDU + \gamma_1 EXP + \gamma_2 EXP^2 + \varepsilon$$

Where Y is the logarithmic value of individual hourly earnings in the labor market, EDU is the indicator variable for their higher education attainment. β is the parameter of interest, capturing the return to higher education. Obtaining a college degree instead of a high school (and another subsequent) qualification is expected to increase individual hourly earnings by about β percent, *ceteris paribus*. EXP stands for individual experience. ε represents the disturbance term.

[53] They adopt a Marginal Treatment Effect (MTE) concept, defined as "the average return to schooling for persons indifferent to going on to schooling at different levels of unobservable factors that determine schooling choices."

The Mincer equation is then extended as follows:

$$Y = \alpha + \beta EDU + \gamma_1 EXP + \gamma_2 EXP^2 + \gamma X + \varepsilon$$

Where X consists of individual observables, which should be exogenous, i.e., unaffected by schoolings. Ideal candidates are pre-treatment characteristics and time-invariant factors (Blundell et al., 2004) such as gender, ethnicity and region.[54] In contrast, sector indicators are 'bad' controls as one benefit of education is that it facilitates sectorial choice, thus including them in the earnings equation might underestimate returns to schooling (Heckman & Li, 2004).

The inclusion of control(s) for family background such as parental education and/or parental income is not necessarily desired. On the one hand, literature acknowledges that these variables could also proxy individual ability (Heckman & Li, 2004; Doan, 2011 among others).[55] Although family background determining personal ability might not be a very convincing statement, including this (these) variable(s) probably reduces bias due to the unobserved. On the other hand, the potential upward bias caused by the unobserved ability and family background may be partially offset by the attenuation measurement error bias (Blundell et al., 2004). Therefore, OLS estimates without ability and background controls might be sufficiently close to the real impact of higher education on individual earnings. In addition, if the information on family background is only available for an unrepresentative subsample, the regression without family background controls should be preferred.

Since in LFS 2007, earnings are not observed for everyone, those with available information on earnings might not be representative of the whole labor force. If self-selection bias is significant, the obtained estimates are valid only to the income-earners and cannot be generalized to the whole population. The Heckman procedure (*Heckit* method) is therefore applied to account for the sample selection issue and to assess the stability of the estimates. The Heckman selection model relies on identifying variables which strongly affect one's probability of getting some kind of income but not their income level. In our model, marital status and whether the individual is head of their household are chosen as excluded variables. Married household heads, bearing more financial pressure than the others, are more likely to work; but the fact that they are married and heads of household hardly influences their earnings.

[54] Region is included in the model, keeping in mind that migration potentially 'endogenizes' this variable. Differently speaking, one's current living location is not necessarily their hometown, and probably affected by their education level.

[55] Family information such as parental education is often used to control for unmeasured ability or as an instrument for children's education (Ashenfelter & Zimmerman, 1997; Card, 1995; Conneely & Uusitalo, 1997; Heckman & Li, 2004; Griliches, 1979, as cited in Doan, 2011).

4.2.2. Instrumental Variable Estimation

The causality could be identified by exploiting some 'exogenous' variation in education by using an excluded instrument. A good instrument must be significantly correlated with individual schooling, but uncorrelated with the residuals of the regression equation. In other word, it affects individual earnings only through education.

In the above equation, the instrument could be expressed as a z variable, with $cov(z, EDU) \neq 0$ but $cov(z, \varepsilon) = 0$. The IV estimate could be obtained through the Two Stage Least Squares (2SLS) procedure. In the first stage, EDU is regressed on all other exogenous variables (i.e., EXP, EXP^2 and X), including the excluded instrument z using OLS. This stage would produce the predicted value of education, denoted as $\widehat{EDU}$, which is assumed to be non-correlated with the error term. In the second stage, Y is regressed on $\widehat{EDU}$ and all other exogenous variables, yielding an estimate of the return to education which is unbiased in principle.

The first potential candidate is month of birth. On the one hand, month of birth could be considered naturally exogenous. On the other hand, it may affect individual schooling via several channels (Crawford et al., 2007; Angrist & Krueger, 1991). In Vietnam, school years start in every September and children start going to school in the year when they turn six. Children born toward the end of a year are therefore less mature and less prepared than their peers who are born toward the beginning of a year. Consequently, the latter might outperform and finally achieve higher qualifications.

However, there may be a tendency that parents of the former keep their child at home till the next school year (i.e., when the child turns seven) in order to overcome the aforementioned disadvantage. Meanwhile, children born at the beginning of a year are more likely to join the previous cohort, i.e., they start going to school at the age of five. In this case, those who are born toward the end of a year are more likely to be prepared, thus possibly attaining better educational outcome.

Actually, our data are in favor of the second argument. The first-stage regression shows that month of birth is positively correlated with the probability of completing higher education at five percent level of significance. However, there are some reasons to question the validity of month of birth as an IV for higher education in the earnings equation. First, month of birth might not be as exogenous as expected. Grenet (2010) shows that the birth seasonality is strongly associated with socioeconomic status of the mother. For instance, teachers and professors in France are much more likely to give birth in April, when they can benefit from a 10-week maternity leave that ends right at the beginning of the summer holiday. Second, although month of birth has been found to affect educational attainment, it is questionable that the effect can last as long as to the university level. In fact, Grenet (2010) points out that date of birth influences the type rather than the level of qualification held by individuals.

The second IV candidate is parental education. Standard instrument validity tests performed by Arcand et al. (2004) indicate that parental education satisfies the two usual criteria of a good IV. On the one hand, whether parents graduated from college or not could strongly

affect the possibility of children going to college. Children might be better cared and informed since their early age, as well as encouraged by their educated parents to pursue higher education study. On the other hand, there is little possibility that parental education affects children's earnings via other channel than children's education. In our sample, parental education is captured by a dummy indicating whether one's father has completed higher education or not.

4.2.3. Heterogeneous returns to higher education

As analyzed in the literature review section, the rates of return to education actually vary across the population. In the first step, separate regressions for different sub-samples which may have different returns are performed. Besides, interactions between higher education dummy and several explanatory variables would reveal which individual characteristics are more associated with higher returns to higher education. However, when comparing returns to higher education between college graduates and non-graduates, the most meaningful interaction is between higher education and the propensity to complete it (Heckman et al., 2006). Thus, in a further step, the heterogeneity of the returns to higher education is investigated with two other estimation methods, namely Quantile Regression and Heterogeneous Treatment Effect.

(i) Quantile Regression Estimation

Individual income not only reflects one's current living standard but also gives hints to their social and family background. It would be meaningful to investigate how returns to higher education differ across income groups. A quantile regression model measures the earnings differentials induced by higher education at various quantiles of the log earnings distribution:

$$q_{\varrho}(Y) = \alpha(\varrho) + \beta(\varrho)EDU + \gamma(\varrho)X + \varepsilon, \forall \varrho \in [0,1]$$

Where $q_{\varrho}(Y)$ is the ϱ^{th} conditional logarithm of earnings quantile. The set of coefficients $\gamma(\varrho)$ provides the estimated rates of return to the different covariates at the ϱ^{th} quantile of the log earnings distribution. The coefficients $\beta(\varrho)$ measure the parts of the earnings differentials that are explained by higher education at the various quantiles. The distribution of the error term is left unspecified. Quantile regression provides robust estimates, particularly for mis-specification errors related to non-normality and heteroscedasticity. Another advantage of quantile regression concerns the identification of distributional effects, allowing to test two opposite hypotheses, namely, *sticky floor* and *glass ceiling*. A *sticky floor* effect is at work when the wage gap between those with tertiary education and those without widens at the lower tiers of the earnings distribution, while *glass ceiling* describes the reverse tendency.

(ii) Heterogeneous Treatment Effect Analysis

To assess whether population heterogeneity in the propensity to complete tertiary education is related to heterogeneity in returns to college, a propensity score-based heterogeneous treatment effect analysis (so called *hte* algorithm) is conducted (Jann et al., 2010; Xie et al., 2012). Propensity score is defined as the probability of assignment to the treated group (college degree) given covariates X:

$$P = p(d_i = 1|X)$$

Specifically, the analysis examines if the estimated return to tertiary education is positively or negatively correlated with the estimated propensity to attend college. The *hte* algorithm includes four steps as follows:

(1) Estimation of propensity score by using a *Probit* model.
(2) Construction of balanced propensity score strata where there are no significant differences in the average values of covariates and the propensity score between the treatment and control groups.
(3) Estimation of strata-specific average treatment effects.
(4) Estimation of the trend of treatment effects across propensity score strata.

5. Average returns to higher education: OLS and IV estimates

Table 5 presents the returns to higher education in Vietnam estimated at mean, following different regression models. In general, the OLS estimates (columns (1), (2) and (3)) with different sets of controls are approximately equal, though it seems that the more controls are added, the lower the estimates. In the standard Mincer equation, completing higher education after high school could raise individual earnings by 58.9 percent on average, *ceteris paribus*. This is not much below Doan (2011)'s result at 68 percent. When additional demographic characteristics are controlled for, the estimated return slightly declines to 55.6 percent and 54.3 percent in the models without and with family background controls, respectively. This suggests that the omitted variables in the standard Mincer equations affect one's chance of obtaining a higher education degree and their earnings in the same directions (i.e., upward bias). The negligible difference between models (OLS 1) and (OLS 2) might imply that parental income and education have already been largely reflected through other demographic information. If this is the case, the former model is possibly preferred thanks to its larger sample size, thus more precise estimates.

By contrast, the three IV estimation models yield dispersed results. Using month of birth as the excluded instrument, models (IV 1) fails to detect any significant impact of higher education on individual earnings. Although the reported coefficient is sufficiently close to the OLS estimates, the corresponding standard error is too high, which is likely due to the problem of weak instrument. Meanwhile, using father's education or mother's education as IV, models (IV 2) and (IV 3) produce highly significant estimates that are evidently higher than those obtained with OLS estimation (77.5 percent and 67.4 percent, respectively). This could be attributed to the fact that these are the impact of higher education on the 'compliers,' whose college decision was affected by their parent's educational profile.

Table 5 also shows the significance of all the control variables in almost all specifications. More experience enhances individual earnings up to a certain threshold, after that individual earnings diminishes with too long tenure. In the Mincer equation, on average, individual income

reaches the peak at the age of 47, *ceteris paribus*. This supports Mincer (1974)'s view on self-investment in human capital after the completion of schooling.[56]

In the models (OLS 1) and (IV 1), on average, minor ethnic groups earn less than Kinh people by five percent, and females' earnings level is 11.8 percent lower than males', keeping other things equal. These two variables become non-significant in model (OLS 2) with their corresponding standard errors relatively high due to limited sample size. In all specifications, rural residency is negatively correlated with individual earnings at one percent level of significance. Holding other individual characteristics constant, living in South East, Central Highlands and Mekong River Delta appears to be the most associated with higher earnings among all regions.

The positive and significant coefficients associated with father's income (in logarithmic form) and father's education indicate that more favorable family background is correlated with higher income of individuals. A one percentage increase in father's income is associated with a 4.8 percent increase in individual earnings, *ceteris paribus*. Having a father graduating from a college could raise one's earnings by 10.7 percent on average, other factors being held fixed. These results are in line with the conventional idea that family background, if being omitted from the wage/earnings equation, could cause an upward bias in the estimated return to schooling. However, as mentioned above, this bias is too small to overweigh the detrimental impact induced by narrowed sample size, including some coefficients losing precision and significance.

[56] Mincer (1974) argues that: "Though age can be viewed as an inherent depreciation phenomenon in the human capital terminology, the growth of earnings with age is ultimately interpreted in the human capital model as being a consequence of continued net self-investment activities after the completion of schooling. The theory predicts that investments are concentrated at younger ages, but continue at a diminishing rate throughout much of a person's working life. Because of increasing marginal costs, investments are not incurred all at once in a short period; they are staggered over time and decline continuously —both because benefits decline as the payoff period shortens and because opportunity costs are likely to rise with experience. This is true of gross as well as net investments. Since earnings are a return on cumulated net investments, they also rise at a diminishing rate over the working life and decline when net investment becomes negative, as in old age. The typical (logarithmic) working-life earnings profile is therefore concave."

Table 5. Returns to Higher Education in Vietnam. Dependent Variable: Log Hourly Earnings

VARIABLES	(1) **Mincer**	(2) **OLS 1**	(3) **OLS 2**	(4) **IV 1**	(5) **IV 2**	(6) **IV 3**
Higher education	**0.589*** **	**0.556*** **	**0.543*** **	**0.494**	**0.775*** **	**0.674*** **
	(0.006)	(0.006)	(0.036)	(0.835)	(0.081)	(0.104)
Age	0.094***	0.091***	0.118***	0.090***	0.068***	0.060***
	(0.002)	(0.002)	(0.042)	(0.023)	(0.012)	(0.010)
Age squared	-0.001***	-0.001***	-0.002**	-0.001***	-0.001***	-0.001***
	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)
Ethnic minority		-0.050***	0.067	-0.050***	-0.023	-0.139***
		(0.012)	(0.073)	(0.012)	(0.040)	(0.038)
Female		-0.118***	-0.041	-0.118***	-0.110***	-0.107***
		(0.006)	(0.037)	(0.007)	(0.022)	(0.020)
Rural		-0.327***	-0.118***	-0.337**	-0.192***	-0.323***
		(0.008)	(0.040)	(0.140)	(0.027)	(0.031)
Region (Reference: Red River Delta)						
Northern Midlands and Mountain Areas		0.035***	-0.157**	0.033**	-0.085**	-0.050*
		(0.009)	(0.061)	(0.017)	(0.033)	(0.029)
Northern Central and Central Coastal Areas		0.020**	0.205***	0.021	0.049*	-0.033
		(0.008)	(0.048)	(0.021)	(0.027)	(0.027)
Central Highlands		0.260***	0.331***	0.259***	0.237***	0.210***
		(0.011)	(0.082)	(0.013)	(0.046)	(0.043)
South East		0.378***	0.498***	0.378***	0.386***	0.325***
		(0.011)	(0.063)	(0.018)	(0.034)	(0.033)
Mekong River Delta		0.138***	0.191***	0.139***	0.127***	0.063**
		(0.010)	(0.053)	(0.019)	(0.030)	(0.029)
Father's income			0.048**			
			(0.019)			
Father's education			0.107**			
			(0.050)			
Constant	-0.245***	-0.124***	-0.781	-0.076	0.092	0.404*
	(0.044)	(0.043)	(0.660)	(0.792)	(0.217)	(0.215)
Observations	85,555	85,552	4,761	85,113	10,427	9,222
R-squared	0.116	0.156	0.096	0.155	0.096	0.126

*Standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1; Source: LFS 2007, author's calculation. Note: IV 1: Month of birth; IV 2: Father's education; IV 3: Mother's education.*

6. Heterogeneity of returns to higher education in Vietnam

6.1. OLS estimates

To assess the variation of return to higher education across Vietnamese population, the first step is to run separate regressions on different sub-populations (**Table 6** and **Table 7**). As expected, the return to college education is remarkably heterogeneous across different groups. In the OLS model without controlling for father's income and education, consistent with previous literature (Moock et al., 2003; Nguyen, 2014), females appear to benefit more from tertiary schoolings than do males, at 59.3 percent and 52.1 percent, respectively. The rates of return also differ across ethnic groups and geographical area. On average, minor ethnic individuals would increase their earnings by up to 65.7 percent if they obtain a higher education qualification, compared to only 54.8 percent for Kinh people. Similarly, the impact of attending college seems to be much larger for rural residents (69.7 percent) than for urban ones (53.4 percent). The same patterns are also observed in the OLS model with family background controls, except that in this case, the impact of higher education appears to be larger on males than on females but the difference is not significant.

The return to higher education for informal workers[57] is examined separately to test Moock et al. (2003)'s prediction, that productivity may have a higher payoff in a less tightly regulated informal sector. Surprisingly, it is fairly low (26-27 percent). In fact, the informal workers in this analysis are characterized by relatively low earnings (average hourly income of 6.27 thousand dongs, compared to 9.15 thousand dongs for the whole sample), precarious working condition (78 percent having *'no contract,'* 43 percent having outdoor or mobile premise) and small size of production unit (81 percent having three workers or less). College graduates, often considered as highly qualified workers in the economy, engage in such an employment possibly because some of them could not find a satisfactory formal job. Therefore, university degree might make little difference in this case. Nevertheless, it is possible that informal workers find it easier to hide their real income than do formal counterparts.

In the second step to investigate the heterogeneity of rates of return to higher education, college dummy is allowed to interact with pre-treatment demographic factors, namely '*minor ethnicity,*' '*female*' and '*rural.*' Consistent with the above results, all three interaction terms are positive and highly significant (column (9), **Table 6**), indicating that the return to higher education is larger for these disadvantaged individuals. When family background controls are added (column (9), **Table 7**), the first two interaction terms are no longer significant. However, this is likely due to small sample size. Interestingly, when the interaction terms are included, both OLS models bring the same estimate of return to college education, at 49 percent.

[57] Informal workers include all jobs in the informal sector plus informal jobs in the formal sector (Cling et al., 2010).

Table 6. OLS Estimates of Returns to Education in Vietnam for Different Sub-populations, without Family Background Controls

VARIABLES	(1) **All**	(2) **Males**	(3) **Females**	(4) **Kinh**	(5) **Ethnic minority**	(6) **Urban**	(7) **Rural**	(8) **Informal workers**	(9) **Interaction terms**
Higher education	**0.556*** **	**0.521*** **	**0.593*** **	**0.548*** **	**0.657*** **	**0.534*** **	**0.697*** **	**0.270*** **	**0.490*** **
	(0.006)	(0.008)	(0.009)	(0.006)	(0.024)	(0.006)	(0.020)	(0.022)	(0.009)
Age	0.091***	0.091***	0.093***	0.089***	0.109***	0.094***	0.078***	0.069***	0.091***
	(0.002)	(0.003)	(0.004)	(0.002)	(0.008)	(0.002)	(0.006)	(0.004)	(0.002)
Age squared	-0.001***	-0.001***	-0.001***	-0.001***	-0.001***	-0.001***	-0.001***	-0.001***	-0.001***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Ethnic minority	-0.050***	-0.100***	0.001			-0.025**	-0.023	-0.250***	-0.088***
	(0.012)	(0.016)	(0.017)			(0.013)	(0.031)	(0.027)	(0.014)
Female	-0.118***			-0.130***	0.032	-0.118***	-0.132***	-0.244***	-0.146***
	(0.006)			(0.006)	(0.023)	(0.006)	(0.017)	(0.013)	(0.007)
Rural	-0.327***	-0.328***	-0.330***	-0.313***	-0.435***			-0.330***	-0.376***
	(0.008)	(0.010)	(0.012)	(0.008)	(0.027)			(0.015)	(0.009)
Region (Reference: Red River Delta)									
Northern Midlands and Mountain Areas	0.035***	-0.045***	0.113***	0.038***	-0.052	0.026***	-0.035	-0.055***	0.034***
	(0.009)	(0.012)	(0.013)	(0.009)	(0.103)	(0.009)	(0.029)	(0.020)	(0.009)
Northern Central and Central Coastal Areas	0.020**	-0.009	0.051***	0.020**	-0.004	-0.009	0.141***	0.066***	0.019**
	(0.008)	(0.011)	(0.012)	(0.008)	(0.113)	(0.009)	(0.023)	(0.018)	(0.008)
Central Highlands	0.260***	0.221***	0.302***	0.259***	0.299***	0.197***	0.608***	0.514***	0.259***
	(0.011)	(0.015)	(0.017)	(0.012)	(0.116)	(0.012)	(0.035)	(0.025)	(0.011)
South East	0.378***	0.358***	0.398***	0.378***	0.288**	0.299***	0.711***	0.634***	0.376***
	(0.011)	(0.014)	(0.016)	(0.011)	(0.119)	(0.011)	(0.029)	(0.025)	(0.011)
Mekong River Delta	0.138***	0.115***	0.162***	0.140***	0.026	0.064***	0.479***	0.175***	0.135***
	(0.010)	(0.013)	(0.015)	(0.010)	(0.109)	(0.010)	(0.027)	(0.022)	(0.010)
Edu*Ethnic									0.120***
									(0.023)
Edu*Female									0.072***
									(0.012)
Edu*Rural									0.199***
									(0.018)
Constant	-0.124***	-0.077	-0.322***	-0.082*	-0.563***	-0.160***	-0.255**	0.146	-0.102**
	(0.043)	(0.055)	(0.073)	(0.045)	(0.197)	(0.046)	(0.114)	(0.091)	(0.043)
Observations	85,552	45,355	40,197	79,360	6,192	72,172	13,380	30,167	85,552
R-squared	0.156	0.158	0.152	0.152	0.197	0.129	0.170	0.082	0.158

*Standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1; Source: LFS 2007, author's calculation.*

Table 7. OLS Estimates of Returns to Education in Vietnam for Different Sub-populations, with Family Background Controls

VARIABLES	(1) **All**	(2) **Males**	(3) **Females**	(4) **Kinh**	(5) **Ethnic minority**	(6) **Urban**	(7) **Rural**	(8) **Informal workers**	(9) **Interaction terms**
Higher education	**0.543*** **	**0.555*** **	**0.521*** **	**0.537*** **	**0.645*** **	**0.503*** **	**0.614*** **	**0.261** **	**0.488*** **
	(0.036)	(0.048)	(0.050)	(0.037)	(0.152)	(0.040)	(0.079)	(0.132)	(0.050)
Age	0.118***	0.136***	0.083	0.113***	0.215	0.078*	0.227**	0.106	0.118***
	(0.042)	(0.053)	(0.064)	(0.042)	(0.209)	(0.045)	(0.097)	(0.107)	(0.042)
Age squared	-0.002**	-0.002**	-0.001	-0.002**	-0.003	-0.001	-0.004**	-0.002	-0.002**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.003)	(0.001)	(0.002)	(0.002)	(0.001)
Ethnic minority	0.067	0.101	0.030			0.074	0.339**	0.060	0.007
	(0.073)	(0.096)	(0.103)			(0.084)	(0.155)	(0.182)	(0.089)
Female	-0.041			-0.040	-0.041	-0.024	-0.060	-0.109	-0.051
	(0.037)			(0.038)	(0.148)	(0.041)	(0.082)	(0.102)	(0.050)
Rural	-0.118***	-0.143***	-0.074	-0.111***	-0.210			-0.201**	-0.166***
	(0.040)	(0.051)	(0.060)	(0.041)	(0.153)			(0.097)	(0.048)
Region (Reference: Red River Delta)									
Northern Midlands and Mountain Areas	-0.157**	-0.192**	-0.062	-0.174***	-0.089	-0.051	-0.616***	-0.409***	-0.154**
	(0.061)	(0.077)	(0.097)	(0.065)	(0.438)	(0.067)	(0.144)	(0.152)	(0.061)
Northern Central and Central Coastal Areas	0.205***	0.238***	0.146**	0.211***	0.218	0.160***	0.320***	0.520***	0.204***
	(0.048)	(0.062)	(0.074)	(0.049)	(0.498)	(0.055)	(0.098)	(0.119)	(0.048)
Central Highlands	0.331***	0.412***	0.180	0.347***	0.275	0.316***	0.343	0.748***	0.333***
	(0.082)	(0.108)	(0.114)	(0.083)	(0.534)	(0.085)	(0.236)	(0.204)	(0.082)
South East	0.498***	0.585***	0.365***	0.517***	0.192	0.426***	0.666***	0.898***	0.496***
	(0.063)	(0.086)	(0.084)	(0.063)	(0.516)	(0.071)	(0.131)	(0.179)	(0.063)
Mekong River Delta	0.191***	0.228***	0.115	0.202***	0.057	0.121**	0.348***	0.185	0.186***
	(0.053)	(0.069)	(0.077)	(0.053)	(0.467)	(0.059)	(0.112)	(0.137)	(0.053)
Father's income	0.048**	-0.015	0.181***	0.028	0.228***	0.042**	0.057	0.002	0.048**
	(0.019)	(0.025)	(0.027)	(0.020)	(0.065)	(0.020)	(0.047)	(0.048)	(0.019)
Father's education	0.107**	0.120*	0.084	0.119**	-0.006	0.155***	-0.197	-0.103	0.118**
	(0.050)	(0.067)	(0.069)	(0.052)	(0.228)	(0.052)	(0.166)	(0.180)	(0.051)
Edu*Ethnic									0.175
									(0.140)
Edu*Female									0.020
									(0.073)
Edu*Rural									0.148*
									(0.082)
Constant	-0.781	-0.976	-0.452	-0.660	-2.653	-0.172	-2.462	-0.752	-0.749
	(0.660)	(0.837)	(1.010)	(0.675)	(3.268)	(0.720)	(1.525)	(1.698)	(0.660)
Observations	4,761	3,222	1,539	4,432	329	3,470	1,291	1,645	4,761
R-squared	0.096	0.087	0.150	0.093	0.164	0.082	0.132	0.053	0.097

*Standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1; Source: LFS 2007, author's calculation.*

In the third step, we focus on investigating the returns to higher education for females in different ages. According to Mincer (1974), the impact of age on earnings reflects a post-schooling net self-investment in human capital or productivity changes due to inherent biological and psychological maturation. In case of women, it is important to additionally consider the influence of fertility and maternity leave. The returns to tertiary education for females in different lifetime periods are presented in **Table 8**.

Unsurprisingly, women from 41 to 55 years old have the highest income premium from university qualification (60.9 percent), but it bottoms at just 26 percent for those above 55. While the former are considered in their career 'heyday,' the latter are usually retired and thus presumably benefitting from a pension which is lower than their previous wage rate and losing all extra earnings apart from wages. Indeed, coefficients associated with 'age' for the third group are no longer significant. This result is in line with a fixed pension attributed to the retirees regardless of their age.

Women in usual fertility age (from 25 to 40 years old) benefit just a little less from college education than the female average (57.9 percent versus 59.3 percent). However, this does not mean that fertility has no significant impact on returns to higher education. Indeed, these results are insufficient to isolate the impact of fertility, since other factors are also at work. Moreover, it should be kept in mind that these regressions exclude inactive women without any income sources, who possibly left the labor market due to maternal duties.

Table 8. Returns to Higher Education for Females in Difference Ages

VARIABLES	(1) **25-40**	(2) **41-55**	(3) **>=56**
Higher education	0.579***	0.609***	0.260***
	(0.011)	(0.014)	(0.067)
Age	0.110***	0.339***	-0.007
	(0.019)	(0.042)	(0.153)
Age squared	-0.001***	-0.004***	-0.000
	(0.000)	(0.000)	(0.001)
Ethnic minority	-0.006	0.022	-0.142
	(0.022)	(0.027)	(0.129)
Rural	-0.286***	-0.388***	-0.415***
	(0.015)	(0.021)	(0.082)
Region (Reference: Red River Delta)			
Northern Midlands and Mountain Areas	0.087***	0.162***	-0.112
	(0.016)	(0.020)	(0.091)
Northern Central and Central Coastal Areas	0.019	0.095***	0.061
	(0.016)	(0.020)	(0.079)
Central Highlands	0.306***	0.299***	0.247
	(0.022)	(0.028)	(0.150)
South East	0.383***	0.385***	0.709***
	(0.020)	(0.027)	(0.111)
Mekong River Delta	0.124***	0.189***	0.394***
	(0.020)	(0.024)	(0.101)
Constant	-0.478	-6.355***	1.948
	(0.296)	(0.984)	(4.787)
Observations	23,487	15,573	1,137
R-squared	0.141	0.156	0.091

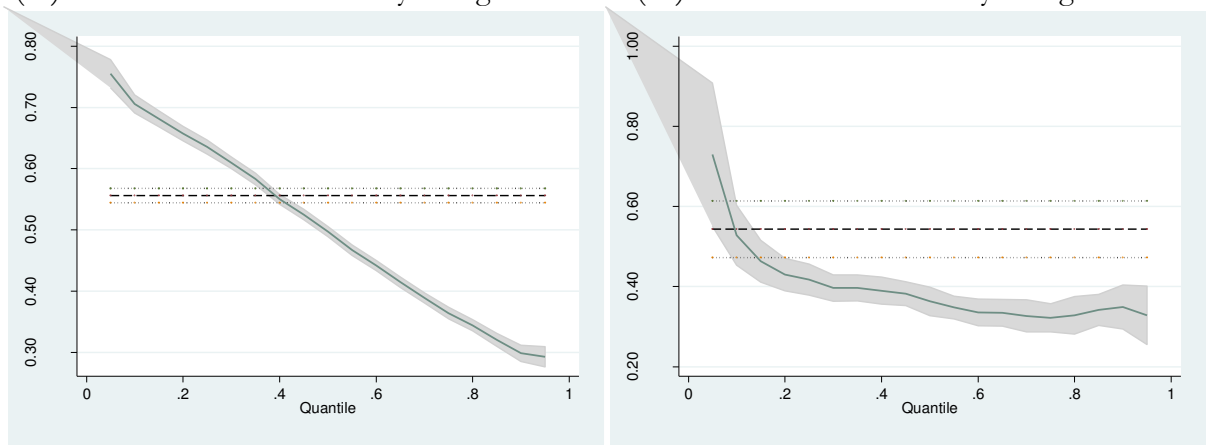
*Standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1; Source: LFS 2007, author's calculation.*

6.2. Quantile Regressions

Another approach toward the heterogeneity of returns to higher education relates to individuals' position in the earnings distribution. The Quantile Regression examining the variation of such returns along various earnings quantiles is presented in **Figure 1** and Table 9.

Figure 1. Returns to Higher Education in Vietnam along the Earnings Distribution

(1a) Without controls for family background (1b) With controls for family background



Source: LFS 2007 (GSO), author's calculation.

It is evident that the earnings premium associated with college education in Vietnam is larger for the individuals in the lower tiers of the earnings distribution than for those in the upper tiers. Such a *sticky floor* effect appears more salient in the model without family background controls: obtaining higher education might raise individual earnings by 70.6 percent at the 10th percentile of the earnings distribution, compared to only 29.9 percent at the 90th percentile, *ceteris paribus*. This illustrates that tertiary education could be more profitable for the disadvantaged workers who are in the bottom of the earnings distribution. When family background is included, the *sticky floor* effect remains at work, but the gap drops steeply in the lowest quantiles and the confidence intervals enlarge. The earnings premium associated with higher education gradually declines from 52.8 percent at the 10th quantile to 32.3 percent at the 75th quantile of the earnings ladder, before slightly rising to 34.9 percent at the 90th quantile.

In both models, the signs of other covariates stay the same across reported quantiles except for regional dummies. However, the magnitudes of some explanatory variables substantially vary along the earnings distribution. To be specific, experience (age), ethnicity and rural/urban area seem to matter the most to the low-income earners. For example, in the left hand side model, living in the countryside instead of cities could lower hourly earnings level by 53.2 percent for those located in the 10th earnings percentile, while only reduces that of those in the 90th percentile by 14.6 percent, *ceteris paribus*. This is explainable as the impact of these background factors is likely to fade away when people already reach very high income levels. Interestingly, father's income is not significant in the full sample regression, but turns into significant in five reported quantiles. This illustrates one important advantage of quantile regression, which is to detect the significant effect of variables which appear insignificant in the regression at mean.

The sticky floor effect observed in Vietnam is consistent with Wang (2013)'s findings in China. Recall that Wang (2013) considers earnings percentile as the level of individual ability, implying that abler individuals receive lower marginal returns to schooling than do less able individuals. In this paper's setting, that 'ability' could be interpreted as individuals' probability of obtaining higher education, which is not only the consequence of their innate ability but also affected by their given resources. Accordingly, this result signifies that those who are less likely to attend college tend to benefit more from it in terms of earnings.

Wang (2013)'s explanations in China's context seem greatly relevant in Vietnam, due to two countries' similarities. First, if education and abilities are truly substitutes, they will be negatively correlated. Indeed, less variation in earnings gap from around the 10^{th} percentile in the extended model could be related to this hypothesis of substitutability between education and propensity to pursue education, partially reflected through family background. Once family resources are controlled for, the earnings gap approaches the 'pure' higher education premium, which is not that variant as before.

Second, Vietnamese schools are comparable to Chinese counterparts with regard to an over-emphasis on exam-taking techniques while essential skills are sometimes ignored. As workers often have fewer skills than required, education matters more among low-ability ones. This is in line with the aforementioned result that going to college is much more beneficial for those who are at the bottom of the earnings distribution, even when the impact of family background has been isolated.

Third, Vietnam is also characterized by a restrictive mobility system as China (*ho khau*, i.e., household registration). Less able or disadvantaged individuals tend to be less mobile, taking children in mountainous areas as an example. A tertiary study, however, could allow them to migrate to developed cities for better-paid jobs. As a consequence, larger return to higher education in lower earnings quantiles implies that education could abate earnings inequality while raising individual earnings. Differently speaking, education-enhancing policies could simultaneously achieve efficiency and equity.

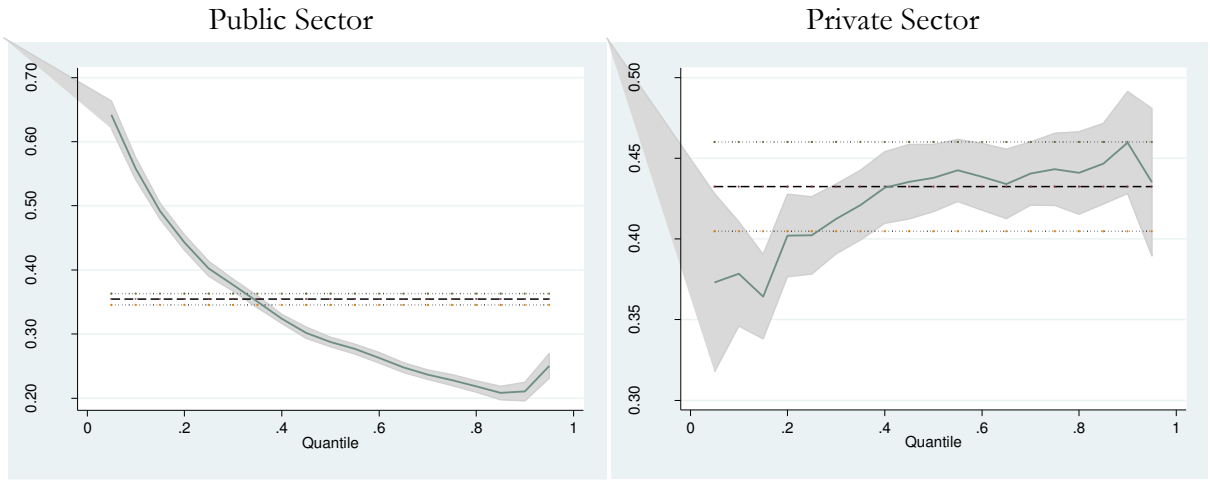
Another interpretation stated in Di Gropello (2006) relates to labor market rigidities. The author argues that in the labor markets driven mainly by market forces, the complementarity of skills is reflected in pay, and the pattern of increasing returns with earnings is likely to occur. Whereas, in the economies where market forces are depressed by labor market rigidities and state intervention, returns tend to decrease with earnings, particularly in the public sector. It is certainly the case in Vietnam. To further investigate this argument, separate quantile regressions for public and private sectors are implemented (**Figure 2**). The results do validate this hypothesis. While a sticky floor effect is very evident in the public sector, there is no significant distribution effect in the private sector, with the point estimate even increasing with earnings.

Table 9. Returns to Higher Education in Vietnam - Quantile Regressions

	Without controls for family background						With controls for family background					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
VARIABLES	**All**	**QR .10**	**QR .25**	**QR .50**	**QR .75**	**QR .90**	**All**	**QR .10**	**QR .25**	**QR .50**	**QR .75**	**QR .90**
Higher education	**0.556*** **	**0.706*** **	**0.635*** **	**0.497*** **	**0.364*** **	**0.299*** **	**0.543*** **	**0.528*** **	**0.417*** **	**0.363*** **	**0.322*** **	**0.349*** **
	(0.005)	(0.008)	(0.005)	(0.004)	(0.004)	(0.006)	(0.032)	(0.040)	(0.020)	(0.016)	(0.016)	(0.026)
Age	0.091***	0.103***	0.092***	0.081***	0.066***	0.056***	0.118***	0.195***	0.147***	0.093***	0.070***	0.068***
	(0.002)	(0.004)	(0.002)	(0.002)	(0.002)	(0.003)	(0.042)	(0.049)	(0.033)	(0.018)	(0.023)	(0.023)
Age squared	-0.001***	-0.001***	-0.001***	-0.001***	-0.001***	-0.001***	-0.002**	-0.003***	-0.002***	-0.001***	-0.001**	-0.001**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)
Ethnic minority	-0.050***	-0.259***	-0.086***	0.011	0.029***	0.030***	0.067	0.004	-0.065	0.008	0.069*	0.093**
	(0.013)	(0.023)	(0.016)	(0.010)	(0.008)	(0.011)	(0.079)	(0.058)	(0.060)	(0.044)	(0.037)	(0.044)
Female	-0.118***	-0.118***	-0.109***	-0.091***	-0.086***	-0.115***	-0.041	-0.039	-0.073***	-0.070***	-0.097***	-0.136***
	(0.006)	(0.007)	(0.006)	(0.004)	(0.004)	(0.006)	(0.033)	(0.038)	(0.019)	(0.016)	(0.016)	(0.026)
Rural	-0.327***	-0.532***	-0.409***	-0.245***	-0.176***	-0.146***	-0.118***	-0.199***	-0.075***	-0.027	-0.030	-0.092***
	(0.009)	(0.012)	(0.011)	(0.008)	(0.006)	(0.009)	(0.043)	(0.044)	(0.027)	(0.020)	(0.019)	(0.026)
Region (Reference: Red River Delta)												
Northern Midlands and	0.035***	0.049***	0.031***	0.037***	0.017**	-0.025**	-0.157*	-0.019	-0.056	-0.061**	-0.129***	-0.216***
Mountain Areas	(0.010)	(0.012)	(0.009)	(0.007)	(0.007)	(0.010)	(0.082)	(0.087)	(0.042)	(0.028)	(0.026)	(0.048)
Northern Central and	0.020**	0.060***	0.016**	-0.010*	-0.055***	-0.102***	0.205***	0.234***	0.058**	0.012	-0.045**	-0.076*
Central Coastal Areas	(0.009)	(0.011)	(0.008)	(0.006)	(0.006)	(0.009)	(0.052)	(0.057)	(0.028)	(0.020)	(0.021)	(0.043)
Central Highlands	0.260***	0.280***	0.221***	0.162***	0.125***	0.093***	0.331***	0.215**	0.146***	0.048	0.040	-0.045
	(0.010)	(0.013)	(0.010)	(0.008)	(0.009)	(0.013)	(0.058)	(0.084)	(0.051)	(0.034)	(0.042)	(0.075)
South East	0.378***	0.358***	0.291***	0.253***	0.244***	0.281***	0.498***	0.382***	0.234***	0.197***	0.243***	0.198***
	(0.009)	(0.015)	(0.009)	(0.008)	(0.010)	(0.015)	(0.055)	(0.074)	(0.034)	(0.032)	(0.037)	(0.050)
Mekong River Delta	0.138***	0.154***	0.108***	0.092***	0.068***	0.080***	0.191***	0.215***	0.072**	0.047*	0.076***	0.039
	(0.011)	(0.014)	(0.010)	(0.007)	(0.008)	(0.013)	(0.062)	(0.062)	(0.033)	(0.027)	(0.025)	(0.040)
Father's income							0.048	0.114***	0.083***	0.080***	0.069***	0.079***
							(0.030)	(0.024)	(0.010)	(0.009)	(0.009)	(0.014)
Father's education							0.107**	0.024	0.047**	0.015	0.052**	0.045
							(0.050)	(0.048)	(0.023)	(0.020)	(0.024)	(0.037)
Constant	-0.124**	-0.887***	-0.369***	0.137***	0.745***	1.279***	-0.781	-2.548***	-1.212**	-0.103	0.607*	0.989***
	(0.051)	(0.069)	(0.048)	(0.038)	(0.036)	(0.052)	(0.663)	(0.778)	(0.510)	(0.291)	(0.364)	(0.377)
Observations	85,552	85,552	85,552	85,552	85,552	85,552	4,761	4,761	4,761	4,761	4,761	4,761
R-squared	0.156						0.096					

*Standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1; Source: LFS 2007, author's calculation.*

Figure 2. Quantile Regression without Family Background Controls: Public versus Private Sectors



Source: LFS 2007 (GSO), author's calculation.

6.3. Heterogeneous Treatment Effects

Previous estimation has discovered the heterogeneous returns to higher education among different subpopulations in Vietnam. On the one hand, whether disadvantaged individuals could actually better profit from college needs to be verified through a more refined estimation method. On the other hand, if in reality, these people are also less likely to obtain higher education, it becomes even more meaningful. In this section, a *'Stratification-Multilevel method'* (SM method, Xie et al. 2012) is applied to elucidate those hypotheses.

The first step in the SM method is to estimate the propensity to complete higher education by using a Probit model. **Table 10** presents estimation results of four models explaining higher education dummy as a function of individual characteristics. Model 1 (full specification) includes age, ethnicity, gender, rural/urban area, region, father's income and education as explanatory variables. Model (2) excludes family background controls. Model (3) further drops regional dummies. Model (4) keeps regional dummies but omit ethnicity, as these two information might be highly correlated with each other.

In three out of four models, older generations are less likely to complete tertiary education than the younger ones, but the impact magnitude of age is just minimal (-0.5 percentage point). Gender female appears to be associated with a higher opportunity of entering universities in the full model. In all specifications, the probability of finishing higher education is much lower for rural residents than for urban citizens (-48 percentage points in models (2) and (4)). Taking Red River Delta as the reference region, individuals from Northern Midlands and Mountain Areas are less likely to graduate from college, other things equal. In contrast, those from Northern Central and Central Coastal Areas, South East and Mekong River Delta are more capable to obtain higher education.

Meanwhile, *'ethnic minority'* is not statistically significant in models (1) and (2) where regional dummies are present. This is possibly because regional variables have already captured a great

amount of information concerning ethnicity. Indeed, as region dummies are excluded from the regression (Model (3)), the coefficient of ethnicity turns into significant (-3.9 percentage points). In model (4), when *'ethnic minority'* is dropped and region dummies are kept, the estimated results are very close to model (2)'s estimates.

Family background is highly correlated with individual propensity of going to college (model (1)). On average, a one percent increase in father's income could raise individual probability of obtaining higher education by 12.8 percentage points, *ceteris paribus*. Impressively, children of college graduates are 78.9 percentage points more likely to go to college than these of non-graduates according to model (1)'s results.

Table 10. Who Are More Likely to Go to College? Probit Estimations.

	(1)	(2)	(3)	(4)
VARIABLES	Probit 1	Probit 2	Probit 3	Probit 4
Age	0.004	-0.005***	-0.005***	-0.005***
	(0.005)	(0.000)	(0.000)	(0.000)
Ethnic minority	-0.085	0.013	-0.039**	
	(0.082)	(0.019)	(0.017)	
Female	0.228***	-0.001	-0.005	-0.001
	(0.041)	(0.009)	(0.009)	(0.009)
Rural	-0.215***	-0.480***	-0.471***	-0.479***
	(0.045)	(0.013)	(0.013)	(0.013)
Region (Reference: Red River Delta)				
Northern Midlands and Mountain Areas	-0.020	-0.056***		-0.053***
	(0.069)	(0.014)		(0.013)
Northern Central and Central Coastal Areas	0.185***	0.061***		0.061***
	(0.054)	(0.013)		(0.013)
Central Highlands	0.021	0.004		0.004
	(0.090)	(0.018)		(0.018)
South East	0.093	0.045***		0.045***
	(0.070)	(0.017)		(0.017)
Mekong River Delta	0.097	0.049***		0.049***
	(0.059)	(0.015)		(0.015)
Father's income	0.128***			
	(0.023)			
Father's education	0.789***			
	(0.057)			
Constant	-0.759***	-0.125***	-0.109***	-0.125***
	(0.160)	(0.022)	(0.020)	(0.022)
Observations	4,771	85,589	85,589	85,592

*Standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1; Source: LFS 2007, author's calculation.*

To sum up, there exists evidence that individuals from disadvantageous background (ethnic minority, females, rural residents, remote geographical region, poorer and less educated parents) are possibly less inclined to complete higher education. Given that there should be no significant difference in terms of inborn qualities needed for pursuing education among ethnic groups, genders, regions and different family backgrounds; the aforementioned result probably reflects excessive constraints that prevent the under-privileged from receiving college education. Besides the liquidity constraint which is the most acknowledged in the literature and in the policies, there could be other infrastructural and cultural hurdles. Minor ethnic groups, for examples, often live in remote mountainous areas with poor traffic conditions, sparse schools, little information, early marriage custom, where people lack a recognition of education's benefit. Likewise, females are less encouraged to pursue higher education than males in the Oriental traditional culture.

The second step in the SM method constructs balanced propensity score strata, i.e. there are no significant differences in the average values of covariates and the propensity score between the treatment and control groups. In the next step, average treatment effects are estimated for each stratum. The final step estimates the trend of treatment effects across propensity score strata. **Table 11** and **Figure 3** present the returns to higher education across propensity strata, with and without family background control(s).

In **Table 11**, smaller strata order (e.g., strata 1, 2, 3) indicates lower propensity of completing college of individuals within these strata.[58] The estimated results reveal consistently positive and significant rates of return to higher education in almost all propensity score strata. However, the estimated returns vary greatly across different strata. Tertiary education is associated with a 50 percent increase in hourly earnings for those who are most likely to obtain it in model (1) and 30.8 percent in model (2), *ceteris paribus*. Meanwhile, a college degree might roughly double individual income for those who are the most unlikely to attend college (stratum 1 in both models). In general, there exists a negative linear trend between propensity score strata and the impact of higher education, which is significant at one percent level. However, adding family background controls reduces the extent of such a negative trend (from -0.042 to -0.035).[59]

Figure 3 visualizes the pattern of estimated returns to higher education across various propensity score strata. The horizontal axis presents propensity score strata, the higher it is, the more likely to attain a college education. The vertical axis projects the point estimates of the returns to higher education in corresponding propensity score strata and their confidence intervals. In general, **Figure 3** illustrates that, those who are less likely to obtain higher education would benefit more from it. This result supports the Negative Selection Hypothesis proposed in the sociological literature and confirmed by Brand and Xie (2010).

[58] The number of strata presented in the regression table and figure is produced automatically by the "hte" command in STATA.

[59] Brand and Xie (2010) also find a declining trend of return to higher education across propensity score strata (i.e., negative selection), using the full set of covariates. When they restrict models to a limited set of covariates, omitting ability and academics, social-psychological variables, and religion, evidence for positive selection is observed. The authors argue that this is because these additional variables "exert greater power explaining college completion in higher propensity score strata than in lower propensity score strata."

Table 11. Returns to Higher Education by Propensity Strata

Without controls for family background

Strata	Coef.	Std. Err.	z	P>|z|	95% Conf. Interval	
1	0.968	0.091	10.600	0.000	0.789	1.147
2	0.794	0.033	24.140	0.000	0.729	0.858
3	0.941	0.095	9.910	0.000	0.755	1.127
4	0.561	0.026	21.300	0.000	0.509	0.613
6	0.617	0.035	17.710	0.000	0.549	0.685
7	0.559	0.009	62.660	0.000	0.541	0.576
8	0.502	0.008	59.120	0.000	0.485	0.519
Linear trend						
Slope	**-0.042**	0.004	-9.890	0.000	-0.050	-0.033
Constant	0.840	0.030	27.630	0.000	0.780	0.899

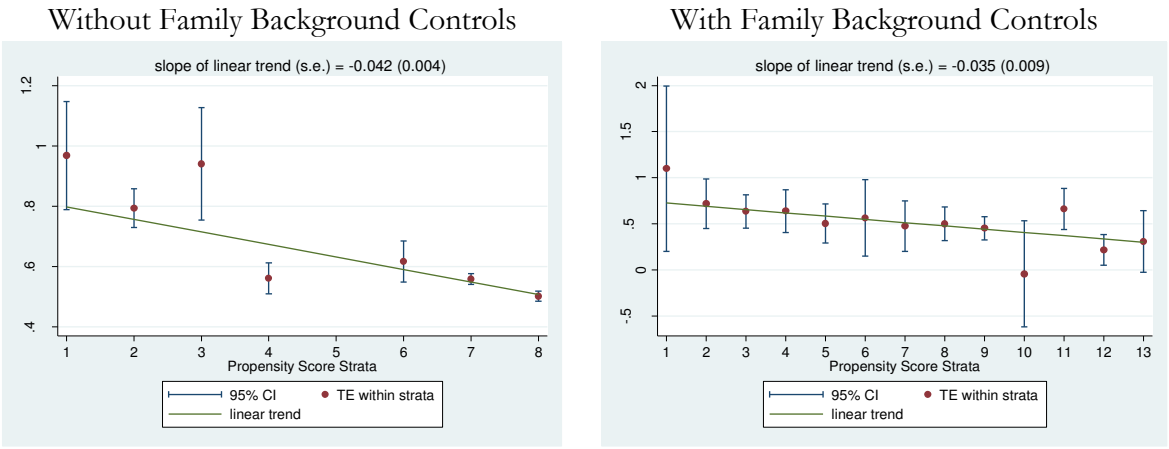
Note: Propensity score is constructed based on *ethnicity, gender* and *rural/urban area*. Within-strata treatment effects is estimated using *age* and *age squared, ethnicity, gender, rural/urban area* and *region*. The balancing property is satisfied. Number of observation: 104,206.

With controls for family background

Strata	Coef.	Std. Err.	z	P>|z|	95% Conf. Interval	
1	1.099	0.457	2.410	0.016	0.203	1.995
2	0.720	0.137	5.250	0.000	0.451	0.989
3	0.636	0.092	6.890	0.000	0.455	0.817
4	0.639	0.118	5.410	0.000	0.408	0.870
5	0.505	0.109	4.640	0.000	0.292	0.718
6	0.564	0.212	2.660	0.008	0.149	0.979
7	0.475	0.140	3.400	0.001	0.202	0.749
8	0.502	0.093	5.410	0.000	0.320	0.684
9	0.452	0.064	7.080	0.000	0.327	0.578
10	-0.042	0.293	-0.140	0.886	-0.617	0.533
11	0.662	0.114	5.830	0.000	0.439	0.884
12	0.218	0.085	2.560	0.010	0.051	0.384
13	0.308	0.171	1.810	0.071	-0.026	0.642
Linear trend						
Slope	**-0.035**	0.009	-3.720	0.000	-0.054	-0.017
Constant	0.761	0.079	9.620	0.000	0.606	0.917

Note: Propensity score is constructed based on *age, ethnicity, gender, rural/urban area, region* and *father's education*. Within-strata treatment effects is estimated using *age* and *age squared*, *ethnicity, gender, rural/urban area, region, father's income* and *father's education*. The balancing property is not satisfied. Number of observation: 12,167. *Source: LFS 2007, author's calculation.*

Figure 3. Returns to Higher Education in Vietnam across Propensity Score Strata



Source: LFS 2007, author's calculation.

According to this branch of literature, mechanisms affecting college decision may differ by social background. Higher education is less attached to economic gain for some people from socially advantaged backgrounds than it is for those from disadvantageous social backgrounds. For the former, going to college may be a cultural norm; whereas, for the latter, due to resource constraints, college decision must be made based on careful economic considerations (Beattie, 2002; Boudon, 1974; Smith & Powell, 1990). This argument is particularly relevant in Vietnam, where the traditional Confucian culture reserves an honorable position to the highly educated. Although the tuition fees in Vietnam's public universities are relatively low, expensive costs of living in big cities like Hanoi and HCMC might hinder students from poor, rural, remote areas to pursue higher studies.

In addition, Brand and Xie (2010) argue that the relatively large return to higher education of low-propensity university graduates could be explained by the fact that their social position is marked with substantial disadvantage. Without a college degree, these individuals have limited human and social capital that disables their labor market prospects. In the contrary, in the absence of university qualification, individuals from more advantaged social backgrounds could still benefit from their superior resources. For the former, university may be used as a means for economic mobility, whereas for the latter, college is a cultural norm rather than an economic decision. Brand and Xie (2010) also emphasize that the negative selection phenomenon does not appear because low-propensity college goers earn more than do high-propensity counterparts. Instead, the pattern arises "because low-propensity non-college goers earn so little." In Vietnam's context, this argument is likely to hold. Vietnam's labor market, especially in the public sector, has been facing a lack of transparency, the importance of personal relations and corruption. Without a college degree, workers from lower social classes with no personal connections are unlikely to attain good positions with higher income.

Moreover, larger benefit from higher education for people from lower socioeconomic backgrounds could also relates to the fact that, to complete tertiary schooling, they have to overcome much more obstacles than those from more advantaged backgrounds. Di Gropello (2006) notes that demand side constraints are ubiquitous, particularly for lower-income

households. First, due to uncertainty about labor market returns to higher education and inadequate information on schooling options, they may underinvest in education and make short-sighted schooling decisions. Second, school-related financial burden might be unbearable to some poor families, which induces them to work in parallel with the study and increases their dropout rate. Moreover, lower-background students may often graduate from lower-quality high schools (e.g., most of the best high schools in Vietnam are located in the cities). For those who decide to take the university entrance examination, to pass it is another obstacle. To sum up, those who could overcome their deprived conditions to successfully complete tertiary schooling are undoubtedly competent, and thus having created a huge gap between them and the low-propensity high school leavers.

Another possible interpretation of the negative selection effect is that higher education might be scarcer among the lower-background community than among the higher-background one. If one's working environment is surrounded by those from the similar background (e.g., rural workers work in rural areas), the university qualification owned by the disadvantaged might be further valued and thus, bringing larger earnings premium. Remind that the estimated returns in our analyses have already incorporated the effect of government's policies, possibly including the subsidies to workers under poor working conditions such as teachers and doctors in mountainous regions. For example, the decree number 61/2006/ND-CP states that teachers and education managers working in areas with unfavorable socioeconomic conditions can receive a subsidy of 70 percent their current wage rate.[60] Likewise, the document number 3766/TC-GD-YT-DTMN-TCCQ issued by Binh Thuan province also mentions that students coming from some remote mountainous districts could benefit from a 100 percent subsidy if they return to their hometown to work as teachers and doctors after graduation.[61] Consequently, larger return to higher education estimated for the disadvantaged may not only capture their actual gain from college, but also reflect the bias induced by state intervention. However, as these subsidies are only targeted at a limited occupational group in specific situations, the overall pattern of negative selection is unlikely to be misled.

[60] Source: http://www.moj.gov.vn/vbpq/lists/vn%20bn%20php%20lut/view_detail.aspx?itemid=15943.

[61] Source: http://thuvienphapluat.vn/van-ban/Tai-chinh-nha-nuoc/Quyet-dinh-03-2002-QD-UBBT-tro-cap-giao-vien-can-bo-y-te-hoc-sinh-mien-nui-vung-cao-hai-dao-Binh-Thuan-160011.aspx.

7. Robustness checks

This section examines the stability of the estimated returns to higher education in Vietnam across different estimation models, how they change in response to modifications in relevant samples and the credibility of some results. Recall that the OLS regressions of individual earnings on their college attainment and a set of covariates produce estimates of 54-59 percent (**Table 5**). These analyses are applied on all income earners of at least 25 years old, having completed their student life, and restricted to those who have finished at least 12 years of general education. First, Heckman selection model is implemented to take into account the potential difference between income earners and the rest of the population. Second, the analysis is restricted to wage workers only, excluding all employers, self-employed, family workers and apprentices. Third, the restriction on those who have completed high school is relaxed. Results are summarized in **Table 12**.

The first two columns present Heckman estimates of returns to higher education with two different sets of covariate, i.e. without and with family background controls. Although the first stage estimates indicate the presence of selectivity (lambdas are negative and significant), the main equation estimates are virtually equal to those of the OLS models, at 55.3 percent and 54.5 percent. Therefore, it is probable to rely on OLS and subsequent estimates without correcting for sample selection.

The next two columns show the impact of higher education on hourly earnings for wage workers in Vietnam. Wage workers appear to benefit less from college in terms of earnings than the sample average. In the model without family background covariates, on average, obtaining higher education could raise a wage worker's earnings by 41.4 percent, *ceteris paribus*. When father's income and father's education are included in the equation, the higher education premium drops to 33.4 percent. This could be related to the fact that the income of salaried workers is more rigid than other types of workers such as self-employed and family workers.

Columns (5) and (6) describe the earnings differentials between college graduates and the rest of the sample, including those who have not finished high school yet. In principle, these estimates should be larger than the estimates in the previous section, for they capture not only the impact of higher education but also the cumulated effect of lower schooling levels. For example, the earnings gap between a university graduate and someone who has just obtained primary education would reflect the joint effect of lower secondary school, high school and university education. This is actually the case. On average, college goers earn up to 78-79 percent more than non-goers, *ceteris paribus*.

Table 12. Returns to Higher Education in Vietnam – Robustness Checks

VARIABLES	(1) **Heckman 1**	(2) **Heckman 2**	(3) **WW 1**	(4) **WW 2**	(5) **HE 1**	(6) **HE 2**
Higher education	**0.553*** **	**0.545*** **	**0.414*** **	**0.334*** **	**0.778*** **	**0.798*** **
	(0.006)	(0.036)	(0.004)	(0.016)	(0.006)	(0.042)
Age	0.037***	0.034	0.077***	0.086***	0.058***	0.062**
	(0.004)	(0.059)	(0.002)	(0.019)	(0.001)	(0.031)
Age squared	-0.000***	-0.000	-0.001***	-0.001***	-0.001***	-0.001**
	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)
Ethnic minority	-0.033***	0.258**	0.042***	0.008	-0.307***	-0.246***
	(0.012)	(0.117)	(0.008)	(0.033)	(0.006)	(0.061)
Female	-0.060***	-0.037	-0.035***	-0.085***	-0.223***	-0.147***
	(0.007)	(0.039)	(0.004)	(0.016)	(0.004)	(0.034)
Rural	-0.366***	-0.070	-0.131***	-0.020	-0.358***	-0.238***
	(0.009)	(0.048)	(0.006)	(0.018)	(0.004)	(0.032)
Region (Reference: Red River Delta)						
Northern Midlands and Mountain Areas	0.018**	-0.091	0.044***	-0.068**	0.013**	-0.356***
	(0.009)	(0.072)	(0.006)	(0.028)	(0.006)	(0.060)
Northern Central and Central Coastal Areas	0.013	0.252***	-0.025***	-0.010	0.112***	0.406***
	(0.009)	(0.056)	(0.006)	(0.022)	(0.006)	(0.046)
Central Highlands	0.236***	0.437***	0.110***	0.071*	0.512***	0.744***
	(0.012)	(0.100)	(0.009)	(0.037)	(0.007)	(0.075)
South East	0.400***	0.521***	0.182***	0.193***	0.578***	0.823***
	(0.011)	(0.068)	(0.008)	(0.027)	(0.007)	(0.059)
Mekong River Delta	0.170***	0.289***	0.040***	0.035	0.265***	0.335***
	(0.010)	(0.073)	(0.007)	(0.024)	(0.006)	(0.047)
Father's income		0.094***		0.077***		0.011
		(0.029)		(0.009)		(0.017)
Father's education		-0.017		0.036*		0.226***
		(0.079)		(0.022)		(0.062)
Constant	0.913***	0.700	0.225***	0.022	0.332***	0.063
	(0.087)	(0.982)	(0.034)	(0.300)	(0.024)	(0.502)
Observations	104,120	5,722	55,512	3,767	267,365	10,663
R-squared			0.230	0.208	0.184	0.106
lambda	-0.545***	-0.879**				
	(0.039)	(0.408)				

*Standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1; Source: LFS 2007, author's calculation.*

The fourth robustness check concerns the disparities between the estimates without and with family background controls. Apart from different model specifications, whether they are due to different samples needs to be checked. By construction, the subsample with available information on family background consists of the adults who still lived in the same household with their parents. **Table 13** presents a comparison between the subsample with family background information and the rest. In general, it indicates that the former was generally younger, more dominated by males, more likely to stay in rural areas and to work as a wage workers than the latter. The rate of higher education completion was higher for the former group, yet their average hourly earnings were lower. There is no significant differences in terms of ethnic composition and institutional sectors (i.e., informality and public sector) between the two groups. Overall, these findings suggest that the estimation results with family background controls might not be representative of the whole population of interest. Thus, probably we had better rely on the regressions without controlling for parental income and education.

Table 13. Sub-sample with Available Information on Family Background: Is There Sample Selection?

Variable	Without information		With information		Difference
	Mean	Std. Err.	Mean	Std. Err.	
Age	40.391	0.033	28.246	0.054	12.145***
Ethnic minority	0.073	0.001	0.069	0.004	0.003
Female	0.479	0.002	0.323	0.007	0.155***
Rural	0.150	0.001	0.272	0.006	-0.123***
Wage workers	0.641	0.002	0.791	0.006	-0.151***
Informal employment	0.353	0.002	0.347	0.007	0.006
Public sector	0.497	0.002	0.499	0.007	-0.002
Hourly earnings	9.250	0.033	7.506	0.090	1.744***
Higher education	0.355	0.002	0.418	0.007	-0.064***

Source: LFS 2007, author's calculation.

Next, it is important to examine if the 'negative selection' phenomenon suggested by the heterogeneous treatment effect analysis as well as the *sticky floor* effect observed in quantile regressions truly imply that those who are less able to accomplish college study tend to gain more from it. Recall that this statement is only valid for those who have finished twelve years of general education. In some cases, completion of high school might be an indication of exceptional ability or resources, such as for ethnic minority. On the supply side, there are very few high schools in remote mountainous areas which are characterized by poor transport and living conditions. On the demand side, minor ethnic families confront budget constraint, information inadequacy and lack of interest in pursuing education due to social norms. Therefore, those with certain unfavorable endowments yet persisting till the end of high school probably possess valuable personal traits and other favorable conditions. To illustrate this conjecture, we examine if minor ethnic high school graduates are significantly 'superior' to their non-graduate counterparts in other aspects (**Table 14**).

Table 14. Are Minor Ethnic High School Graduates Truly Exceptional?

Variable	Not completed high school			Completed high school			Difference
	Mean	Std. Err.	Obs.	Mean	Std. Err.	Obs.	
Age	40.680	0.068	23946	39.225	0.120	6197	1.455***
Female	0.446	0.003	23946	0.495	0.006	6197	-0.049***
Rural	0.661	0.003	23946	0.249	0.005	6197	0.412***
Father's income	4.174	0.345	602	7.096	0.428	332	-2.922***
Mother's income	3.951	0.426	393	9.068	1.452	178	-5.117***
Father completing higher education	0.004	0.002	1797	0.109	0.004	749	-0.105***
Mother completing higher education	0.001	0.001	2334	0.039	0.007	684	-0.038***
Wage workers	0.119	0.002	23940	0.663	0.006	6195	-0.544***
Informal employment	0.931	0.002	23946	0.350	0.006	6197	0.581***
Public sector	0.042	0.001	23946	0.589	0.006	6197	-0.547***
Hourly earnings	3.785	0.035	23946	8.469	0.082	6197	-4.684***

Source: LFS 2007 (GSO), author's calculation.

As expected, the ethnic minority having completed high school were much more likely to benefit other supportive conditions than those who have not. While two third of the non-graduate group lived in rural areas, only a quarter of the graduate group did so. The family background of the latter group was greatly better. For instance, whilst merely 0.4 percent of the high school non-graduates had a father with a higher education degree, the rate was nearly eleven percent for the high school graduates. It is, however, too soon to conclude that the negative selection hypothesis does not hold because the disadvantaged in our restricted sample were advantageous in some other ways. In fact, when only Kinh individuals are taken into account, similar patterns arise. Kinh people who could complete twelve years of general education also had a noticeably better family and social background than those who could not. Based on **Table 14** and **Table 15**, it is hard to tell if the gap between non-graduates and graduates is greater for the ethnic minority than for the Kinh. The sample selectivity induced by including only high school graduates may affect all subpopulations in the same manner. Our main conclusion thus stays unchanged.

Table 15. High School Non-graduates versus Graduates: Kinh People Only

Variable	Not completed high school			Completed high school			Difference
	Mean	Std. Err.	Obs.	Mean	Std. Err.	Obs.	
Age	42.829	0.028	157876	39.751	0.034	79425	3.077***
Female	0.507	0.001	157876	0.468	0.002	79425	0.494***
Rural	0.428	0.001	157876	0.149	0.001	79425	0.279***
Father's income	6.083	0.125	5324	8.370	0.148	4448	-2.287***
Mother's income	4.698	0.087	3325	7.094	0.277	2338	-2.396***
Father completing higher education	0.011	0.001	11681	0.160	0.004	9692	-0.148***
Mother completing higher education	0.004	0.001	13439	0.064	0.003	8547	-0.060***
Wage workers	0.235	0.001	157799	0.648	0.002	79368	-0.412***
Informal employment	0.871	0.001	157876	0.353	0.002	79425	0.518***
Public sector	0.049	0.001	157876	0.490	0.002	79425	-0.441***
Hourly earnings	5.758	0.019	157876	9.207	0.033	79425	-3.448***

Source: LFS 2007 (GSO), author's calculation.

8. Discussion and conclusion

In the current era of economic development and international integration, Vietnam's economy faces a great demand for high skilled manpower. Tertiary education, therefore, has played a central role in national investment policies. Due to budget constraint, the policy-makers should probably pay particular attention to some groups where the effect could be larger, while assuring a certain level of education opportunity equality in the population. Existing studies have provided some information on the variation of returns to higher education across Vietnamese population, yet, the main objectives of these papers are to estimate the average rate of return, or to examine its evolution over time. This is the first research paper that focuses on the heterogeneity of the returns to higher education across Vietnamese population, based on the LFS in 2007.

The overall estimated rate of return to higher education in Vietnam in this paper is large according to international and regional standards.[62] This is possibly because recent reforms and technology-biased transition process in Vietnam have made high skilled workers increasingly better rewarded, as suggested by the literature. More importantly, the estimated results show that the earnings premium associated with higher education is highly heterogeneous. A propensity-score-based Heterogeneous Treatment Effect model has revealed supporting evidence for *Negative Selection Hypothesis*: those who are least likely to go to college tend to benefit the most from it. This is also in line with the *sticky floor* effect observed in Quantile Regression estimation. Tertiary education appears to be more rewarded for individuals in the bottom of the earnings distribution than for those in the upper tail.

This phenomenon could be explained by the fact that higher education decision might be less based on economic consideration for some advantaged people than it is for the under-privileged. For the latter, without a college degree, limited human and social capital hinders their labor market opportunities. Meanwhile, even without such a degree, individuals from high social backgrounds could still benefit from their superior resources. In other word, the negative selection phenomenon does not mean that low-propensity college goers earn more than do high-propensity counterparts. Instead, it is likely because low-propensity non-college goers earn so little. This is particularly true in Vietnam's context, where the labor market in general and the public sector in particular are characterized by non-transparency and weak functioning of market forces. Without a college degree, workers from a disadvantageous background with no personal connections would have little chance to get a 'profitable' job. In addition, unlike high-propensity university graduates, low-propensity counterparts are those who have overcome various constraints, including but not limited to financial burden, information shortage, weak decision making skill and low quality of general education. These obstacles signify a huge gap between the graduates and the non-graduates of the socially lower class. Furthermore, the university qualification owned by the disadvantaged might be relatively rare, thus being further appreciated and bringing larger earnings premium.

The negative selection evidence suggests that investing in the disadvantaged could address the issue of unequal education access on the one hand and is cost effective on the other hand. To be detailed, thanks to higher education, the under-privileged could raise their earnings by a greater

[62] World average: 19%; Asia: 18.2% (Psacharopoulos & Patrinos, 2004).

proportion, thus contributing to the national budget through paying income tax, which compensates part or all of government's initial educational subsidies. The total social gain (i.e., from both the subsidized individuals and the State) could be positive as a result. However, it should be reminded that the analyses in this paper only present private returns to higher education, and it is impossible to firmly conclude about the actual social returns from this setting. Ideally, policy recommendation to the government should be based on social returns. Nevertheless, the prevalence of credit constraint and market imperfection in Vietnam might justify government's intervention in this particular case.

The main finding of this paper motivates policies that encourage the socially disadvantaged individuals to pursue higher education. Yet, to define appropriate policies, the government should start from their own barriers to college and address them in an integrated manner. Following Di Gropello (2006), a combination of information, mentoring and financial incentives can be used to deal with low demand for higher education from lower social background groups. Financial measures include well-targeted grants, loans, and savings schemes. In addition, policies such as bonus grades for ethnic minority and rural habitants in university entrance examination could also encourage the disadvantaged's access to higher education. It should be noted that tertiary education is the direct outcome of twelve years of general schooling; therefore, the coverage and quality issues of primary and secondary education cannot be ignored.

There remain several limits of this paper which could be handled in further research. First, due to data constraint, higher education attainment has been coded as a dummy, ignoring differences in the exact number of years of schooling, majors and quality. The data only allow to distinguish between those who have obtained a college degree and those who have not; whereas, in reality, there could be some individuals who have attended college without achieving a diploma. The time spent in universities might bring useful knowledge, skills and social network, which could affect future earnings. In addition, some majors seem more promising than others, and the employability as well as remuneration of different domains may vary with the macroeconomic situation in the economy. Likewise, recruiters may welcome students from prestigious universities more than those from less known schools. This limit suggests careful context-based considerations while encouraging the underprivileged to attend college. To afford university study, these individuals might have to mobilize important family resources, which would potentially put them in vulnerable situations if they could not find a good job after graduation. For example, the mass creation of banks about ten years ago has led to a 'human resource fever' which has caused huge returns to tertiary education in banking and finance. High school graduates in the ending period of such fever, who had not anticipated the mass bankruptcy of Vietnamese banks soon after, attended banking programs, and finally graduated with poor labor market opportunity. Recently, Vietnamese media have reported that numerous university graduates who are unemployed for several years have to put away their degrees to take a vocational course ('trung cấp nghề').[63]

Second, the set of control variables could have been enriched had more individual information been included in the survey questionnaire such as number of siblings, mental ability,

[63] Reference : "Masters and university graduates putting away their degrees to study vocational programs" (http://vnexpress.net/tin-tuc/giao-duc/thac-si-cu-nhan-cat-bang-di-hoc-trung-cap-3390136.html)

encouragement from teachers to attend university and so on. Third, the impact of migration or geographical shift has been ignored in this analysis. In fact, the geographical variables capturing rural/urban area and regions indicate an individual's current situation rather than their original place. Therefore, some of the so-called urban residents in the analysis have actually migrated from the countryside, whose return to higher education might be even larger than the non-migrant rural residents. For this, the actual negative selection phenomenon could be of an even greater extent.

It is widely believed that Vietnam's enormous potential has not yet been achieved due to the weaknesses of its higher education system. The experience of East and Southeast Asian countries such as South Korea and Taiwan has revealed a close relationship between economic development and higher education. This paper further confirms the positive and large impact of higher education on individual earnings as having been broadly addressed in the literature. It also provides a potential suggestion to the Vietnamese policy makers: investing in the disadvantaged could probably achieve both economic efficiency and equality. It should be kept in mind that the benefit of higher education, however, exceeds monetary values.

Appendix

Table 16. Literature on Returns to Education in Vietnam

Paper	Data	Dependent variable	Methodology	Relevant Results	Note
Stroup and Hargrove (1969)	Rural Income and Expenditure Sample Survey of South Vietnam, 1964	Income: either individual earnings from wages and salaries; or all individual income	OLS	Returns to one year of education: 16.8%.	The range of educational attainment found among the rural Vietnamese is too narrow and the level too low to allow study of differing rates of return to differing educational levels
Gallup (2002)	VLSS 1993 and 1998	Hourly wage	OLS (Mincer and extended)	Returns to one year of education: 2% in 1993; 4%–5% in 1998. The lowest rates of return to schooling are in the regions with the lowest education levels, and the highest rates of return are in the regions with the highest levels of schooling. Women face significant wage discrimination after controlling for their schooling and work experience, but the estimated wage gap fell by half over the period.	
Moock et al. (2003)	VLSS 1992-1993	Earnings of wage labor force	OLS Mincerian	Private returns to higher education: 11% Average private return on another year of schooling: 5%	Establish a benchmark for future work The estimate could be higher if the self-employed were included

				Returns are higher for: females, public sector.	There are factors other than education, for example Communist Party membership, that distort public sector pay
Arcand et al (2004)	VLSS 1992-93 and 1997-98 (panel).	Hourly wage including all reported wages, bonuses, work subsidies and income in kind. 324 wage-earning males.	OLS	OLS: 2.6%; IV: 7% (parental education), 5.1% (HT set of instruments), 6.5% (combination of two IVs)	Y = log hourly wage, including all reported wages, bonuses, work subsidies, and income in kind.
			IV, with Hausman-Taylor matrix of instruments	Standard instrument validity tests suggest that many variables do not satisfy the necessary conditions allowing them to be used as instrumental variables. IV estimates of the returns to education are substantially higher than the corresponding OLS estimate. Hausman-Taylor matrix of instruments, when combined with other instruments, may be a useful way of consistently estimating an average return to education rather than a local average treatment effect	
Liu (2006)	VLSS 1992-93 and 1997-98	Earnings, only wage earners	OLS Mincerian (for 1998 round) Correction for sample selection: Hay's (1979) approach, a generalization of Heckman's approach (for 1993 round)	Returns to one year of education: Men: 6% in 1993; 3.5% in 1998. Women: 4.2% in 1993; 4.8% in 1998	

			Katz–Murphy's demand–supply method	The wage structure has changed in such a way as to favor better-educated people during Vietnam's transition.	
Doan and Gibson (2009)	VLSS 1998 and 2004	Hourly wages including bonuses, allowances and subsidies (both in cash and in-kind)	OLS (Mincerian and extended) Correct for sample selection bias (Heckman) Correct for endogeneous choice of employment sector	Basic Mincerian: 2.7% in 1998; 8.6% in 2004. Extended Mincerian: 1.8% in 1998; 2.5% in 2004. Heckman: 3.8% in 1998; 9.4% in 2004 Control for self-selection into the state sector: 1.6% in 1998; 4.7% in 2004	The trend of increased rates of return appears to be robust to both forms of self-selection considered, that of selection into the wage labor force and selection into a state sector job.
Doan and Gibson (2010)	VLSS 1998, VHLSS 2002, 2004, 2006 and 2008	Hourly wages including bonuses, allowances and subsidies (both in cash and in-kind)	OLS Mincerian Heckman correction for sample selection bias Pooled data with interaction terms	Returns to one year of education: 2.9% in 1998 rising to 9.5% in 2008 Returns to one year of education: 3.8% in 1998 increasing to 10% in 2008 Strong evidence against the hypothesis that the returns to schooling are constant over the period	
Doan (2011)	VHLSS 2008	Hourly wage	OLS, IV and Treatment Effect Model IV used: family background (parental education, assets and share of university and post graduated members in family)	Return to university education: 68% and robust to various estimators Estimated return using IV models with valid instruments is almost the same with that based on the OLS model with family background controls.	
Oostendorp and Quang Hong Doan (2012)	VLSS 1998, VHLSS 2002, 2004 and 2006	Hourly wage	Modified OLS Mincerian, with sample selection correction and employment sector controls	Returns to education increase over time. From 7.4% in 1998 to 13.3% in 2006 (for workers with 15 years of education - college level)	

			Model the sector choice (multinomial logit)	Employment across sectors varies systematically with education. Negative employment effect: Wages tend to be higher in industries with relatively more low skill workers.	Ignores regional shifts in employment (migration). Household composition, marital status and household size variables are included in the employment model but not in the Mincer equations (exclusion restrictions).
			Re-estimate the Mincer Equation each sector separately	The returns to education tend to be the highest in the nontradable and import-substituting sectors. The returns to education increase over time in the non-tradable and import-substituting sector, but stay relatively stable in the export-oriented sector.	
			Unconditional returns to education = wage effect + employment effect	Unconditional returns to education: 15.5% in 1998, 21.7% in 2006 (for college level)	
Diep Phan (2012), published as Phan and Coxhead (2013)	VHLSS 1993 to 2008	Hourly wage (salary plus cash bonuses and kin-kind benefits)	OLS Mincerian	The returns to one year of education gradually increases from 1.9% in 1993 to 5.8% in 2008.	
			Dummies for educational level	1993: a college degree resulted in a log wage only 27.2% higher than for those with no education. The return increased steadily from 1993 through 2008, return to college degrees increased the most (79.2% in 2008).	

			Heckman correction for sample selection bias	Sample selection exists, but no major qualitative differences between OLS and Heckman estimates	Identification variables include the dependency ratio, a household head dummy variable, and non-wage income (non-wage income variables are not available for survey years 1993 and 1998).
			Control for endogeneous selection into state sector jobs, using a treatment model	Without controlling for this, the estimated return seems upward biased, but no qualitative changes.	
Nguyen Dung Tien (2014)	LFS 2012 (panel: 4 quarters)	Hourly earnings	Random effects estimation: Basic and extended Mincer equation, Heckman correction	Returns to education: 4% for the primary level, 24.6% for lower secondary, 54.4% for upper secondary and 68% for college or university (one additional year of university: 4.5%). Estimated returns are higher for: women; public sector; industry and services; subsectors of (i) banking, finance and insurance, (ii) science, technology and professional activities. Having a high education level is more significant in the formal sector.	

Source: Author's synthesis.

Table 17. Vietnam's Labor Force Survey in Different Rounds: A Comparative View

		2007	2009	2010 (2nd round)	2011	2012
Sampling	**Sample size**	173,000 households (HHs)	18,000 households	123,000 households	155,014 households	102,651 households
	Representativeness	Rural/Urban Regions & 64 provinces	Rural/Urban 8 domains (6 regions + Hanoi & HCMC)	Rural/Urban Regions & 63 provinces	Rural/Urban Regions & 63 provinces	Rural/Urban Regions & 63 provinces
	Sampling design	Two-stage stratified sample: - Random selection of enumeration areas 5,768 EAs - Random selection of 30 HH in each EAs) Based on 1999 Population census	Two-stage stratified sample: - Random selection of enumeration areas - Random selection of HH in each EAs Based on 15% census sample of 2009 Population census	Two-stage stratified sample: - Random selection of enumeration areas (3,890) 60 EAs /provinces + 74 EAs in HN & HCMC - Random selection of 32 HHs in each EAs Based on 15% of 2009 Population census (Master sampling frame)	2-2-2 Rotative sampling. Each enumeration area is divided into 2 rotation group, households will be included in 2 adjacent quarters, after that it will be excluded in next 2 adjacent quarters, and will be included in the followed 2 adjacent quarters - Random selection of enumeration areas 60 EAs /provinces (2 samples, 30 urban & 30 rural) + 180 EAs in HN & HCMC - Random selection of 20 HHs in each EAs each month. 80 HH/EA/year Based on 15% of 2009 Population census (Master sampling frame)	Rotative sampling Based on 15% of 2009 Population census (Master sampling frame)

Socio-demographic characteristics	**Education**	9. At present, is [NAME] attending, or has attended or never attended school? 10. What is the highest grade of general school completed by [NAME]? (0-->12)	No information on educational level	9. What is the highest grade of education/training (regular and non-regular) that you have been attended or graduated?	11. What is the highest grade of education/training (regular and non-regular) that you have been attended or graduated?	12. What is the grade of education/training that you are currently attending? (same modalities as 2010)
		11. What is the highest technical qualification completed by [NAME]?	14. What is the highest technical qualification you completed? 15. What was the field of study of the technical/vocational training you completed?			13. What is the highest grade of education/training (from 3 months and over) that you have been attended or graduated? (same modalities as 2010)
	Others*	yes	yes	yes	yes	yes

Source: General Statistics Office. Note: EA = enumeration area. (*) includes ethnic group, marital status, sex, age, relationship to the household head.

References

Angrist, J. & Krueger, A. (1991). Does Compulsory School Attendance Affect Schooling and Earnings?. *The Quarterly Journal of Economics*, *106*(4), 979-1014.

Angrist, J., & Krueger, A. (1999). Empirical strategies in labor economics. *Handbook of labor economics*, *3*, 1277-1366.

Arcand, J., D'Hombres, B. & Gyselinck, P. (2005). Instrument Choice and the Returns to Education: New Evidence from Vietnam. Labor and Demography, EconWPA.

Ashenfelter, O., & Zimmerman, D. (1997). Estimates of the Return to Schooling from Sibling Data: Fathers, Sons and Brothers. *Review of Economics and Statistics, 79*(1), 1-9.

Beattie, I. (2002). Are all Adolescent Econometricians Created Equal? Racial, Class, and Gender Differences in College Enrollment. *Sociology of Education*, *75*, 19-43.

Belzil, C. (2007). The Return to Schooling in Structural Dynamic Models: a Survey. *European Economic Review, 51*(5), 1059-1105.

Blundell, R., Dearden, L., & Sianesi, B. (2004). *Evaluating the Impact of Education on Earnings in the UK: Models, Methods and Results from the NCDS*. Centre for the Economics of Education, London School of Economics and Political Science.

Boudon, R. (1974). *Education, Opportunity and Social Inequality*. New York: John Wiley & Sons.

Brand, J., & Xie, Y. (2010). Who Benefits Most from College?: Evidence for Negative Selection in Heterogeneous Economic Returns to Higher Education. *American Sociological Review*, *75*(2), 273-302.

Card, D. (1999). The causal effect of education on earnings. *Handbook of labor economics*, *3*, 1801-1863.

Card, D., & Krueger, A. (1992). Does School Quality Matter? Returns to Education and the Characteristics of Public Schools in the United States. *Journal of Political Economy*, *100*(1), 1-40.

Carneiro, P., Heckman, J., & Vytlacil, E. (2003). Understanding What Instrumental Variables Estimate: Estimating Marginal and Average Returns to Education. *Processed, University of Chicago, The American Bar Foundation and Stanford University, July*, *19*.

Cling, J. P., Nguyen, T. T. H., Nguyen, H. C., Phan, T. N. T., Razafindrakoto, M., & Roubaud, F. (2010). *The informal sector in Vietnam: a focus on Hanoi and Ho Chi Minh City*. Hanoi: The Gioi Editions.

Coleman, J. (1988). Social Capital in the Creation of Human Capital. *American Journal Of Sociology*, *94*, S95-S120.

Crawford, C., Dearden, L., & Meghir, C. (2007). *When you are born matters: The impact of date of birth on child cognitive outcomes in England*. Centre for the Economics of Education, London School of Economics and Political Science.

Di Gropello, E. (2006). *Meeting the Challenges of Secondary Education in Latin America and East Asia: Improving Efficiency and Resource Mobilization*. Retrieved from http://siteresources.worldbank.org/INTEAPREGTOPEDUCATION/Resources/Meeting-Challenges-of-Secondary-EDU.pdf.

Doan, T. (2011). Labor Market Returns to Higher Education in Vietnam. *SSRN Electronic Journal*, (2011-4).

Doan, T., & Gibson, J. (2012). Return to Education in Vietnam during the Recent Transformation. *International Journal of Education Economics and Development*, *3*(4), 314-329.

Glewwe, P., & Jacoby, H. (2004). Economic Growth and the Demand for Education: is There a Wealth Effect?. *Journal of Development Economics*, *74*(1), 33-51.

Glewwe, P., & Patrinos, H. (1999). The Role of the Private Sector in Education in Vietnam: Evidence from the Vietnam Living Standards Survey. *World Development*, *27*(5), 887-902.

Goldthorpe, J. & Jackson, M. (2008). Education-Based Meritocracy: The Barriers to its Realization. In A. Lareau & D. Conley, *Social Class: How Does it Work?* (pp. 93-117). New York: Russell Sage Foundation.

Grenet, J. (2010). Academic Performance, Educational Trajectories and the Persistence of Date of Birth Effects. *Evidence from France. London School of Economics (mimeo)*.

Hausman, J., & Taylor, W. (1981). Panel Data and Unobservable Individual Effects. *Econometrica*, *49*(6), 1377-1398.

Hayden, M., & Lam, T. (2015). A 2020 vision for higher education in Vietnam. *International Higher Education*, *44*.

Heckman, J., & Li, X. (2004). Selection Bias, Comparative Advantage and Heterogeneous Returns to Education: Evidence from China in 2000. *Pacific Economic Review*, *9*(3), 155-171.

Heckman, J., Urzua, S., & Vytlacil, E. (2006). Understanding instrumental variables in models with essential heterogeneity. *The Review of Economics and Statistics*, *88*(3), 389-432.

Imbens, G., & Angrist, J. (1994). Identification and Estimation of Local Average Treatment Effects. *Econometrica*, *62*(2), 467-475.

Jann, B., Brand, E., & Xie, Y. (2010). *Heterogeneous Treatment Effect Analysis*. Presentation, Berlin.

Kane, T. & Rouse, C. (1995). Labor-Market Returns to Two- and Four-Year College. *American Economic Review*, *85*(3), 600–614.

Liu, A. (2006). Changing wage structure and education in Vietnam, 1992-98. *The Economics of Transition*, *14*(4), 681-706.

Mincer, J. (1974). *Schooling, experience, and earnings*. New York: National Bureau of Economic Research; distributed by Columbia University Press.

Moock, P., Patrinos, H. & Venkataraman, M. (2003). Education and earnings in a transition economy: the case of Vietnam. *Economics of Education Review*, *22*(5), 503-510.

Nguyen, D. T. (2014). An analysis of labour market returns to education in Vietnam: Evidence from the National Labour Force Survey 2012. *Turin School of Development Working Paper*, No. 3, International Training Centre of the ILO, Turin, Italy.

Oostendorp, R. H., & Doan, Q. H. (2013). Have the Returns to Education Really Increased in Vietnam? Wage versus Employment Effect. *Journal of Comparative Economics*, *41* (3), 923-938.

Phan, D., & Coxhead, I. (2013). Long-run Costs of Piecemeal Reform: Wage Inequality and Returns to Education in Vietnam. *Journal of Comparative Economics*, *41*(4), 1106-1122.

Psacharopoulos, G., & Patrinos, H. (2004). Returns to Investment in Education: A Further Update. *Education Economics*, *12*(2), 111-134.

Smith, H., & Powell, B. (1990). Great expectations: Variations in income expectations among college seniors. *Sociology of Education*, *63*, 194-207.

Stroup, R., & Hargrove, M. (1969). Earnings and Education in Rural South Vietnam. *The Journal of Human Resources*, *4*(2), 215-225.

UNESCO. (2014). *Higher Education in Asia: Expanding Out, Expanding Up: the rise of graduate education and university research.* UNESCO Institute for Statistics. Retrieved from http://www.uis.unesco.org/Library/Pages/DocumentMorePage.aspx?docIdValue=809&docIdFld=ID.

Wang, L. (2013). How Does Education Affect the Earnings Distribution in Urban China?. *Oxford Bulletin of Economics and Statistics*, *75*(3), 435-454.

World Bank (1998). Vietnam Development Report 1999: Preparing for Take-offs? Hanoi.

World Bank (2008). Vietnam: Higher Education and Skills for Growth. Retrieved from http://siteresources.worldbank.org/INTEASTASIAPACIFIC/Resources/Vietnam-HEandSkillsforGrowth.pdf.

Xie, Y., Brand, J., & Jann, B. (2012). Estimating Heterogeneous Treatment Effects with Observational Data. *Sociological Methodology*, *42*(1), 314-347.

Chapter 3

Wage Differentials between Temporary and Permanent Workers in Developing Asian Countries

Co-authored with Nguyen Huu Thanh Tam and Nguyen Huu Chi

Abstract

Despite the rise of non-standard jobs in the past several decades, the empirical evidence on wage differentials between temporary and permanent workers is extremely limited in developing Asian countries. This paper is the first research work that systematically examines the temporary-permanent wage gap in selected Asian countries, based on their national Labor Force Survey data. Our estimates indicate the presence of a temporary employment's wage penalty in Indonesia, Pakistan, the Philippines and Vietnam, and contrarily, a wage premium in Cambodia. Moreover, Quantile Regression Estimates show that wage differentials could greatly vary across the wage distribution. Both *sticky floor* and *glass ceiling* effects appear. The wage gap is wider at the bottom of the wage distribution in Vietnam and Pakistan, suggesting that the penalty of being in temporary jobs could be more severe for disadvantaged workers, whereas the *glass ceiling* effect in Indonesia prevents temporary workers from approaching high wages. We argue that country-specific institutional and socioeconomic contexts could possibly explain these different distribution patterns of wage gap, and country-specific measures should be applied to address this issue.

Acknowledgements

This paper was developed from our project on non-standard employment in Asia funded by the International Labor Organization (ILO) in Geneva, the Switzerland. Our commissioned report for this project has been published as the 69th ILO working paper of the Conditions of Work and Employment Series in 2016. We are also grateful to the ILO for providing the data and continuous feedbacks from Mariya Aleksynska and other colleagues in ILO. This paper has been greatly improved thanks to comments received at the 12th International Conference of the Western Economic Association International in Nanyang Technological University, Singapore; at the DIAL seminar; and at the A/D Seminar of Laboratory of Study and Research on Economics, Policies and Social Systems (LEREPS), Toulouse, France.

1. Introduction

The past several decades have witnessed the rise of non-standard forms of employment (NSFE), understood as jobs that do not conform to criteria of standard employment (i.e., full-time and indefinite), in many parts of both the developed and the developing worlds (ILO, 2015a; Serrano et al., 2014 among others). One important category of NSFE is temporary employment, referring to jobs with limited duration, namely fixed-term, project or task-based contracts, seasonal and casual work, including day-laborers (ILO, 2015a). In Europe, for example, the temporary employment rate increased from nine percent in 1987 to 15.2 percent in 2006 (ibid).

Temporary jobs, on the one hand, could provide firms as well as workers with greater adjustment and flexibility. On the other hand, they are often associated with poor job quality, in terms of lower earnings, higher level of labor market insecurity and higher job strain (OECD, 2014). Existing literature has usually raised the question of whether temporary workers being discriminated against in terms of payment or not, the extent and distribution pattern of such discrimination; since most of the theories, such as Human Capital, Efficiency Wage, Labor Market Duality or Insider-outsider models put forward various reasons why the former may receive lower wages, and available empirical evidence also supports this wage penalty. However, there is still a possibility that temporary employees are better remunerated to compensate for other non-wage disadvantages (the theory of Equalizing Difference).

Therefore, empirical studies on the wage gap between temporary and permanent jobs are of great importance for better understanding challenges associated with temporary employment as well as developing policy proposals.[64] Yet, in developing Asian countries, evidence on temporary-permanent wage differential is extremely limited. Most available studies on Asia have been for Japan and the Republic of Korea (Ahn, 2006; Lee & Lee, 2007; Diamond, 2010; Ha & Lee, 2013 among others) or on the formal-informal wage gaps in less-developed nations (Nguyen et al., 2013; Dasgupta et al., 2015 for examples). Moreover, existing evidence on temporary-permanent employment's wage differential mainly relies on simple 'raw' wage gaps, while little is known about the 'pure' wage gaps derived from empirical wage equations. Furthermore, although studies in developed economies suggest that wage penalties highly depend on the relative position in the wage distribution, to the best of our knowledge, such evidence is non-existent in Asian developing countries.

This paper makes an important contribution to the literature as the very first systematic research work examining the wage gap between temporary and permanent workers in developing Asian economies. The analyses are based on the national Labor Force Survey (LFS) data of Cambodia, Indonesia, Pakistan, the Philippines and Vietnam provided by the ILO. Interestingly, evidence for both temporary workers' wage penalty and premium is found. On the one hand, temporary workers suffer from wage penalties of at least two percent in Indonesia, Pakistan, the Philippines and Vietnam; on the other hand, they benefit from a wage premium of five percent in Cambodia. Vietnam and Pakistan are the countries where the wage penalty is the most severe. In addition, Quantile Regression results show that temporary wage differentials highly depend on the

[64] Wage gap: The difference in rates of pay between two different groups of people (Source: Financial Times -ft.com/lexicon).

relative position in the conditional wage distribution, except in the Philippines and Cambodia, where temporary workers encounter similar levels of wage penalty/premium at any positions in the pay ladder. Both *sticky floor* and *glass ceiling* effects occur. The temporary employment's wage gap is wider in the lower tiers of the wage distribution in Vietnam and Pakistan, suggesting that the penalty of being in temporary jobs could be more severe for disadvantaged workers, whereas the *glass ceiling* effect in Indonesia prevents temporary workers from approaching high wages.

The paper is structured as follows. Section 2 provides an overview of theoretical and empirical literature on the wage differentials related to NSFE and temporary jobs, with a primary focus on Asian evidence. Section 3 presents the economic and labor market contexts of the studied countries, data description and empirical strategies. Results derived from econometric models are analyzed in section 4, followed by concluding remarks in the last section.

2. Literature review

2.1. Theoretical background

To facilitate research on temporary employment, an accurate and comparable definition of this concept should be established. According to ILO (2015a), temporary employment refers to jobs with limited duration, namely fixed-term, project or task-based contracts, seasonal and casual work, including day-laborers. Temporary employment is one important category of NSFE, which is understood as jobs that do not conform to criteria of standard employment (i.e., full-time and indefinite). This section aims to provide background theories on the motivations of engaging in NSFE in general and on the temporary-permanent workers' wage gap in particular.

Why workers and firms seek NSFE?

Firms and workers seek NSFE for different purposes. On the demand side, there are three main reasons explaining the use of NSFE: cost advantages, flexibility, and technological changes.

First, NSFE are often associated with lower wage and non-wage costs for enterprises (von Hippel et al., 1997; Nesheim et al., 2007). It is believed that NSFE are the result of too much protection for standard jobs, including rigid labor regulations (Lee & Eyraud, 2008). For example, the amount of taxes and mandatory contributions on labor paid by the business in Vietnam constitute up to 19.2 percent of their commercial profits. In China, the figure reaches 68 percent (Doing Business, 2009). Thus, using NSFE could be a way that employers evade the mandatory contribution burden. Employing workers managed by third parties can also save firms the expenses in screening, administering and supervising workers (Kalleberg et al., 2003). In the context of developing world, these costs might include 'hidden cost' due to corruption and a lack of transparency. **Figure 3** shows a fairly high level of corruption perception in the studied countries, which possibly induces businesses to seek NSFE with less involvement with the authority.

Second, non-standard employment offers firms greater flexibility, numerically or functionally. Debrah and Smith (2002) assert that the expansion of NSFE is driven largely by

demands for flexibility of employers in the context of deregulated labor markets and intense market competition. Temporary work has always existed to cope with seasonal demand (Harrison & Kelley, 1993) and labor supply fluctuations (Ko, 2003). Since the rise of Toyotism with just-in-time production in the latter part of the twentieth century (Edgell, 2012), organizations are able to hire workers for short time periods, allowing them to control the size of their workforce rather rapidly. Moreover, NSFE can help organizations handle specific, typically short-term tasks, which require special skills unavailable in-house (Kalleberg et al., 2003). Flexibility is particularly desired in developing countries whereby various sources of uncertainty, both economic and politic, prevail.

Third, technological changes have facilitated non-standard forms of work. The development of online services has enabled employees to contact with each other and with the employer, regardless of space and time distances. Consequently, administrative and physical attachment to the organization is limited.

On the supply side, NSFE is either a voluntary or involuntary choice of workers. On the one hand, workers voluntarily seek NSFE to accommodate family, school or other obligations, to acquire extra income, skills, experience and network. Part-time jobs, for instance, are often taken by women who are bound with family responsibilities. A further example is in Australia, where the highest level of job satisfaction is recorded for fixed-term workers (Lee & Eyraud, 2008). On the other hand, NSFE could be an involuntary, situational solution of workers when they are unable to obtain standard jobs immediately. For instance, about a half of non-regular workers in Korea declared that being this kind of workers is an involuntary choice because they need money instantly or they cannot find any satisfactory jobs other than this (Lee & Lee, 2007). A number of studies in Asia have found high transition densities of workers from NSFE into regular paid jobs (Esteban-Pretel et al., 2002; Ghose, 2004; Kim & Lee, 2014).

Temporary workers: wage penalty or wage premium?

Given the abovementioned particularities and motivations of NSFE, it is likely that non-standard workers suffer a penalty in terms of working conditions, particularly wages, compared to their standard counterparts. Various theories have been established to explain the underlying reasons for wage differentials between temporary and permanent workers. Although temporary jobs are often considered lower paid, associated with less favorable working conditions than permanent ones, a temporary work' wage premium is yet justified in the theory.

On the one hand, there are numerous theories indicating a temporary workers' wage penalty, including Human Capital theory, Efficiency Wage theory, Labor Market Segmentation theory and Insider-outsider model (Comi & Brasseni, 2012; Pacheco & Cochrane, 2015). The Human Capital theory attributes wage differentials to the amount of 'human capital' that workers obtain. According to the Becker view, the Gardener view and the Schultz/Nelson-Phelps view, human capital is awarded in the labor market because it raises firms' profit.[65] Meanwhile, the Bowles-Gintis view asserts that higher educated workers would be better paid because they are more obedient to the orders and more reliable members of the firm. The Spence view argues that human capital might be rewarded as they reveal other workers' characteristics. When it is unlikely that temporary workers can finally obtain a permanent contract, they tend to invest in, or be

[65] Lectures in Labor Economics – MIT Economics. Source: *economics.mit.edu/files/4689*

provided with, a smaller amount of firm-specific training, and thus receiving lower wages (Booth et al., 2002). Alternatively, if employers have to invest in higher levels of specific training for temporary workers, this additional cost would unavoidably lead to a temporary wage penalty (Pacheco & Cochrane, 2015).

The Efficiency Wage theory asserts that high monitoring costs and uncertain product demand might induce profit maximizing firms to hire both temporary and permanent workers, with the former receiving a lower wage, even if they are homogeneous and perfect substitutes (Rebitzer & Taylor, 1991). Guell (2000) argues that the opportunity of contract renewal may be used as a 'carrot' to incentivize higher productivity from workers instead of wages.

According to Labor Market Segmentation (or Dual Labor Market) theory, the labor market is composed of two segments. The primary segment typically includes permanent and well paid jobs, particularly in large businesses with a stable work environment and various employment benefits, while the secondary segment is characterized by temporary and badly paid jobs under poor working conditions with high turnover rates (Doeringer & Piore, 1971; Reich et al., 1973).

The Insider-outsider model emphasizes the importance of high firing costs due to rigid labor market regulations. Consequently, firms prefer having a buffer stock of workers (the 'outsider') on fixed-term contracts (FTCs) and with limited employment rights, who can easily be laid off in case of adverse shocks. The risk of dismissal is therefore reduced for permanent employees (the 'insider') (Bentolila et al., 1994), who can then bargain for higher wages. In particular, if a union mainly represents the insiders, firms would employ cheaper temporary workers to cope with union wage pressure. Therefore, the presence of workplace unions is also responsible for the contract duality (Salvatori, 2009).

On the other hand, temporary jobs might offer better wages than permanent counterparts to compensate for their less attractive characteristics. According to the theory of Equalizing Differences, or Compensatory Differences, the total monetary and non-monetary advantages or disadvantages among work activities and among workers themselves could be equalized by the observed wage differentials. As a result, workers could possibly receive compensating wage premiums in doing jobs with undesirable non-wage characteristics (Brown, 1980; Rosen, 1987). In the context of temporary-permanent workers' wage differentials, first, the inconsistent nature of some forms of temporary employment such as seasonal jobs implies times of unemployment, thus, temporary workers should be compensated for this. Second, favorable working conditions associated with long-term jobs such as job security, paid leave and other fringe benefits, could lead permanent workers to accept lower wages, as they are compensated in a non-pecuniary form.

In summary, the extent to which temporary workers are remunerated in comparison with permanent workers is still controversial theoretically. Although the majority of existing theories are in favor of a temporary job wage penalty, such a temporary-permanent workers' wage gap deserves thorough examination in the context of Asian developing countries before any conclusion is made.

***Sticky floor* versus *glass ceiling* effects**

The temporary wage penalty might vary across the wage profile of workers. If it is greater in the lower tiers of the wage distribution, a *sticky floor* effect is at work. Reversely, a *glass ceiling* effect means that the wage gap widens in the upper tail of the wage distribution. The literature on such effects are mostly dedicated to gender wage or earnings gaps (Fang & Sakellariou, 2015; Agrawal, 2013; Christofides et al., 2013; Ge et al., 2011; Arulampalam et al., 2007; Albrecht et al., 2003); however, important lessons could be drawn and applied in the context of temporary wage differentials.

The above-mentioned theories of labor market duality and unions, in fact, imply the *sticky floor* effect. Since a majority of temporary workers, who belong to the disadvantaged labor segment, is found in the lower tiers of the wage distribution, temporary workers are likely to face a larger wage penalty at the bottom of the wage distribution (Comi et al., 2012). Likewise, excessive labor protection for permanent workers resulting from unions could possibly widen the wage penalties at the lower end of the wage distribution (Salvatori, 2009)

A *sticky floor* effect could also come from the absence of minimum wage law, and/or labor markets being more flexible and less regulated, leaving new temporary entrants unprotected from serious wage discrimination. Analogous statements could be found in Ge et al. (2011) and Agrawal (2013) in the context of gender wage gap. By contrast, the presence of equal opportunity, anti-discrimination laws and minimum wage law effectively reduces the wage gap in the lower tail of the wage distribution, which possibly bring about a *glass ceiling* effect. If unions also represent temporary workers, their effort to equalize wages, particularly for low-income earners, could have a similar impact as the minimum wage law.

2.2. Empirical evidence on temporary-permanent workers' wage gap

Generally, existing evidence relies on two types of statistics, one is simply 'raw' wage gap, i.e., the ratio (or its inverse) of temporary employment wage to permanent employment wage, and the other is the wage gap derived from empirical wage equations ('pure' wage gap) in which worker's personal and household characteristics are taken into account. Based on both types of statistics, temporary jobs are generally discriminated in terms of remuneration (ILO, 2008 and 2015; Rani, 2012; OECD, 2014; Gash & McGinnity, 2007; Jahn & Pozzoli, 2013 among others).

The empirical evidence on 'pure' wage penalties associated with temporary employment is extremely limited in developing Asian countries. Most of the available evidence has been referred to the informal-formal wage differentials or 'raw' temporary wage gaps.[66] Hassan and Jandoc (2009) find that in the Philippines in 2006, permanent workers earned on average 51 percent higher than casual workers did. In Indonesia, reported wage ratios were respectively 83.3 percent for FTC workers and 73.8 percent for outsourced workers compared to their permanent counterparts (Akatiga Foundation *et al.*, 2010; cited in Serrano et al., 2014). In Bangladesh, the average wage ratio between casual and regular workers was found remarkable, of around 40 percent in 2010 (ILO, 2013). Among Asian developing economies, India is probably the only country with available 'pure' estimates. Based on wage equations constructed for workers in organized

[66] In Asia, the statistics on 'pure' NSFE-SFE wage gap exist only for the OECD members like Japan, Korea.

manufacturing sector in India, Bhandari and Hesmati (2008) find that permanent workers earned on average 45.5 percent more than non-permanent workers after controlling for different individual human capital as well as job related characteristics.

Regarding the distributional pattern of the temporary wage penalties, to the best of our knowledge, no evidence exists for Asian developing countries. Yet, studies in developed economies suggest that such penalties highly depend on the relative position in the wage distribution. Comi et al. (2012) find out a common *sticky floor* effect among nine European countries. Similarly, in Italy, the temporary employment's wage gap is significantly larger at the bottom of the wage profile and is almost absent for high-wage jobs (Bosio, 2014). Another study by Mertens et al. (2007) points out that, despite being both regarded as 'rigid' economies, Germany and Spain had different experiences of fixed-term jobs. While the *sticky floor* effect is observed in West Germany, the wage penalty shows little variation across the wage distribution in Spain.

In explaining the related factors to the permanent-temporary wage gap, decomposition techniques based on wage equations, such as Oaxaca-Blinder, Neuman-Oaxaca or Machado-Mata procedures (Oaxaca, 1973; Blinder, 1973; Neuman & Oaxaca, 2004; Machado & Mata, 2005), are largely applied. Again, this type of analysis has been conducted more in the studies in developed economies such as European countries and New Zealand (Mertens et al., 2007; Comi & Brasseni, 2012; Bosio, 2014; Pacheco & Cochrane, 2015 for examples); whereas, very little evidence can be found in Asia and especially in Asian developing countries. In India, the wage gap decomposition conducted by Bhandari and Hesmati (2006) has shown that, controlling for various observed characteristics, experience plays the most important role in explaining the pay difference between permanent and FTC workers. An explanation for its effect would be that FTC workers cannot stick to a particular job for a long enough period to be able to reap the benefit of significant wage return like that of a permanent worker.

Given the scarcity of estimates of the wage differential and their distribution pattern associated with temporary employment in developing Asian countries, this paper would make a considerable contribution to the literature. On the one hand, we provide systematic and comparable 'pure' estimates of temporary-permanent workers' wage gaps in emerging Asian economies. On the other hand, this is also the very first research work that applies a distributional approach to assess the temporary wage penalties in these countries.

3. Country context, data and methodology

3.1. A brief introduction of the economy and labor market context in selected Asian countries

This section aims at describing some salient features of the economy and labor market in Cambodia, Indonesia, Pakistan, the Philippines and Vietnam, which might be relevant to the findings of this paper. In general, all of the five selected countries are fast-growing developing economies in Southeast and South Asia, despite some fluctuations caused by the recent crises. They all have an expanding labor force with a large informal sector in parallel with the presence of

unions in the formal economy. Yet, they are diverse in terms of political regime, economic situation, population size, labor market regulations and so on. Indonesia emerged as the wealthiest nation among them with a GDP per capital reaching 10,031 dollars in 2014, followed by the Philippines at 6,649 dollars. Cambodia, Pakistan and Vietnam appeared to have the lowest living standards among five selected countries, at 3,113, 4,590 and 5370 dollars, respectively.[67] The information in this section is mainly collected from Serrano et al. (2014), ITUC (2014), ILO (2015a), ILO (2015b), and ILO (2012), which provide synthetic portraits of the aforementioned countries.

Cambodia was one of the world's fastest-growing economies prior to the economic crisis. Its economic reforms in the transition process from central planning to a market-based economy has brought about political stability, unprecedented economic growth and structural change, although little progress on poverty alleviation and rising inequality was observed (ILO, 2012). Cambodia's real growth estimate has reached seven percent in 2015, and the poverty rate was 17.7 percent in 2012.[68] Since the 1990s, Cambodia's labor force and labor force participation rate have been growing, the employment-to-population ratio has increased in recent years and the unemployment rate was notably low and on a downward trend (ILO, 2012). ILO (2012) considers these as "a symptom of the necessity to work than the economy's ability to create decent jobs." Indeed, the income level and social protection is generally low for a majority of the population, with a dominant share of workers in vulnerable (or informal) employment (73 percent, ibid). The Cambodian Labor Law covers all employees and employers who have an employment contract, either written or oral, including casual workers, thus excluding the bulk of informal sector workers who barely have an employment contract (ibid). Minimum wage exists only in the textile, garment and footwear industry, a main driver of the Cambodian economy, which accounted for 16 percent of national GDP and 70-80 percent of Cambodia's exports.[69] This sector is featured by a broad use of FTCs.

Indonesia is one of the fastest-rising economies in Asia and the biggest economy in ASEAN with average annual growth rate of 5.8 percent in the 2004-2012 period (Serrano et al., 2014).[70] Indonesia also has a large and increasing labor force, with a working age population of 174 million in 2012. Indonesian economy is characterized by a large informal sector, simply defined as own-account workers, self-employed workers assisted by temporary member, casual employees and unpaid workers. Accordingly, about 66 percent of Indonesian workers belong to the informal sector. Despite favorable conditions for union registration, Indonesia witnesses a very fragmented labor movement, low union membership and decreasing union density. The Indonesian Manpower Act (Article 51) allows employment contracts to be either written or verbal agreements, and a majority of Indonesian workers have no written proof of employment, which makes them more inclined to abuse. Most outsourced workers are exposed to vulnerability and discrimination

[67] GDP per capita (PPP, U.S dollars). Source: data.worldbank.org

[68] Source: worldbank.org/en/country/cambodia/overview. Accessed on 17 June 2016.

[69] Source: https://www.cleanclothes.org/resources/publications/factsheets/cambodia-factsheet-february-2015.pdf

[70] The information in this paragraph is synthesized from Serrano et al. (2014).

in the labor market: They can be easily dismissed, are not covered by social security or pension benefits, earn wages lower than the minimum level; older workers are discriminated and married or pregnant workers are retrenched. Companies employ various tactics to hinder the organization of workers, such as requiring workers not to join unions or transferring them to another agency to prevent regularization. Observing many cases of violations of labor regulations, the Indonesian trade unions have been jointly undertaking legislative and political action as well as mobilizations to improve the protection of contract, outsourced and other types of NSFE in Indonesia.

Pakistan is one of the least developed countries in Asia, with a growing semi-industrialized economy that depends on manufacturing, agriculture and remittances.[71] International development indicators of Pakistan are amongst the lowest in the world (ILO, 2015b).[72] The average annual growth rate of five percent since 2005 seems insufficient for a fast-growing population (population growth rate was about 1.5 percent per year) and poverty reduction. Since 2008, Pakistan's economy has encountered macroeconomic instability, fiscal contraction, external account weaknesses and rising inflation. Their recovery from the 2005 earthquake was threatened by the adverse situation in the global market and the war on terrorism. The perceived job volatility further pushes downward pressure on wages. Despite improvements in job opportunities as shown in labor force participation rate (32.8 percent, according to LFS 2008-09), most of Pakistani workers lack decent and productive employment. The informal economy, characterized by low paid and poorly protected jobs, constitutes over 70 percent of non-farm jobs in Pakistan and is largely unorganized. The organized representation of employees and employers mainly exists in the formal economy, but remains low.

The Philippines economy has experienced resilient growth in the past decade, and remarkably recovered after the global economic crisis (ILO, 2015c).[73] In parallel, a fast growing labor force (64.4 percent of the working age population in 2014) challenges the creation of sufficient jobs and decent jobs. In 2014, 6.8 percent were unemployed and 38.6 percent of the employed were in vulnerable employment, a proxy for the informal economy in the Philippines. To reduce risks and vulnerabilities in the labor market due to changing skill demand, global crises and natural disasters, in 2012, the Philippines adopted the Social Protection Operational Framework and Strategy, which particularly targets at increasing employment opportunities and improving protection of workers' rights and welfare. Although service or labor contracting is allowed, the Labor Code imposes their conditions to be legitimate, including the rights to the contractor's employees such as safe and healthy working conditions, labor standards (e.g., service incentive leave, rest days, overtime pay, holiday pay, 13th month pay and separation pay), social security and welfare benefits, the right to self-organization, collective bargaining and security of tenure. The union density and collective bargaining coverage in the Philippines followed a downward trend, which could be attributed to the expansion of the informal economy (ILO, 2012 as cited in Serrano et al., 2014). Fortunately, the Filipino trade unions have employed

[71] Source: tradingeconomics.com/pakistan/gdp-growth. Accessed on 17 June 2016.

[72] The information in this paragraph is synthesized from ILO (2015b).

[73] The information in this paragraph is synthesized from ILO (2015c).

complementary strategies at the enterprise, industry and national levels to accord more protection to non-regular workers in spite of existing difficulties (Serrano et al., 2014).

Vietnam is one of the most populated countries in the world with an estimated population of nearly 90 million in 2012 as well as a large and increasing labor force at about 53 million in 2012 (Serrano et al., 2014).[74] After the crises in the 1970s and 1980s, Vietnam has undergone a period of growth, openness and international integration thanks to the economic reform ('Doimoi'). Its annual GDP growth rate ranged between eight percent and 9.5 percent from 1986 to 1997 and, despite the recent global crisis, remained relatively high at around five percent. The number of enterprises in Vietnam increased annually by 21 percent on average over the 2006-2011 period. Yet, micro, small and medium sized businesses remained the most predominant in the economy with a prevailing informal sector. In 2007, nearly half of non-farm jobs in Vietnam were found in the informal sector (Cling et al., 2010), and 75 percent of temporary jobs were also informal. This sector generally provides low-income jobs with precarious working conditions, without any protection such as labor contract, minimum wage, social security and so on (Cling et al., 2014). A majority of Vietnam's labor force is characterized by a lack of skills. The 2011 LFS shows that 84.4 percent of the entire labor force had never attended any technical training. Only 54 percent of workers in foreign invested enterprises were capable to read and understand their labor contract.[75] The unemployment rate in Vietnam was remarkably low, decreasing from 2.9 percent in 2009 to 1.96 percent in 2012. The Vietnam General Confederation of Labor (VGCL) is the sole trade union in Vietnam, gathering wage workers and all legal freelancers who want to participate. Referring to dispatch workers as temporary workers, Tung (2013, as cited in Serrano et al., 2014) highlights that it is difficult for trade unions to include and support these employees as they only sign contracts with the agency, and that if agencies have trade unions, they are unable to protect temporary workers as they do not operate there.

3.2. Data and descriptive statistics

Data and variables

While cross-country statistics for NSFE, including temporary employment, is available for OECD countries (OECD, 2014), there is a scarcity of such evidence on non-standard work in Asia. Provided a various range of country definitions and of measurement methods, NSFE terms in different studies on different countries are unlikely comparable (Tucker, 2002). In emerging economies and developing countries, non-standard work mainly takes the form of informal employment, particularly in Asia (ILO, 2008; Serrano et al., 2014; Rani, 2008).[76] In this paper,

[74] The information in this paragraph is synthesized from Serrano et al. (2014).

[75] USAID/VCCI The Vietnam Provincial Competitiveness Index 2010.

[76] Informal employment is regarded as "workers in very small firms (fewer than five workers), self-employment, unpaid family work and salaried employment without a proper work contract in the formal sector" (ILO, 2008). In this paper, informal employment refers to unprotected jobs, typically lacking social security as seen in formal employment. In reality, some permanent jobs may be informal, while temporary work could be found in formal enterprises.

temporary employment in each country's context is defined following as much as possible the ILO definition as mentioned in section 2.1 to ensure maximum comparability of the results.

This research work relies on national LFS of Vietnam (2007), the Philippines (2009), Pakistan (2008-2009), Indonesia (2007-2008) and Labor Force and Child Labor Survey of Cambodia (2012). Only wage workers are included in our analyses. In practice, questions concerning employment status and types of labor contract are used to define employees and distinguish those having permanent jobs from those who work temporarily. However, these questions are not similarly compiled in the LFS of the selected countries (**Table 1**). Vietnamese, Cambodian, and Pakistani data allow a computation of temporary employment defined as all jobs with impermanent duration. The questionnaires for these countries include a separate question on labor contract providing information to define permanent wage workers ('indefinite-term labor contract' for Vietnam; 'unlimited duration' for Cambodia; 'Permanent/pensionable job' for Pakistan) and temporary wage workers. An ambiguity arises since there are wage workers who do not have a written contract or in other words, they work based on an oral agreement. Cambodia data provide the most detailed information on both types of contractual engagement (labor contract/verbal agreement) and the associated duration, therefore employees can be distributed into four sub-groups, namely 'limited/unspecified duration oral contracts,' 'unlimited oral contracts,' 'limited/unspecified duration written contracts,' and 'unlimited duration written contracts.' Vietnamese and Pakistani questionnaires integrate information on both types of contract and on duration in one unique question. As a consequence, no detailed information can be available to distinguish unlimited and limited employment among those who engaged in oral contracts or did not know anything about the contract/agreement. In these cases, permanent employment is approximately defined as 'unlimited duration written contracts' without taking into account employees engaging in (unrecognizable) unlimited oral contracts.

Filipino and Indonesian surveys can only provide proxies of temporary employment. In the case of the Philippines, the question on the nature of employment enables only to distinguish permanency in terms of jobs rather than of contracts. Indonesian questionnaire includes a one-for-all question providing both workers' employment status and job permanency. Moreover, wage employees can only be classified into two groups, namely, casual and non-casual. Fortunately, in the Indonesia's context, the concept of casual work might approach the entire temporary jobs. Serrano et al. (2014) indicate that FTCs are considered casual jobs, and that FTC and temporary workers might be directly hired for a specific time period by the principal company, or outsourcing some parts of work or services to other companies.

In this paper, hourly wage is chosen as the dependent variable to assure the comparability among studied countries. In fact, while Indonesian survey has a question directly asking about hourly wage, the Filipino and Vietnamese data only provide daily wage and monthly wage,[77] respectively. Therefore, in the Philippines, hourly wage is computed by dividing daily wage by daily working hour. In case of Vietnam, since the data on monthly working hour is unavailable, hourly wage is proxied as follows:

[77] Vietnam LFS 2007 has one question on total monthly earnings, including wage, bonuses and subsidies. In this paper, we use the term 'wage' for comparability with the remaining countries, but it is actually earnings.

$$hwage_i = \frac{7 * mwage_i}{30 * wwhour_i}$$

Where *mwage* and *wwhours* are respectively monthly wage, and weekly working hour of individual *i*.

Particularly, Pakistani survey provides the information on either weekly wages or monthly wages. In other words, for those whose weekly wage is available, their monthly wage rate is absent and vice versa. For workers with available information on weekly wage, their hourly wage is directly computed by dividing their weekly wage by their weekly working hour. For those with monthly wage information, their hourly wage is approximated following the above equation.

Statistical description

In general, temporary work accounted for an important proportion of wage employment in these Asian countries, with great variation across nations. It spanned from 24 percent in the Philippines to 74.6 percent in Pakistan (**Table 2**), which far exceeded OECD's and Japan's temporary employment rates at around 12 percent and 14 percent, respectively (OECD, 2014). The incidence of temporary employment greatly varied across genders, ages, education levels, institutional sectors, industries and occupations in Vietnam, the Philippines, Pakistan and Indonesia, while distributing rather evenly among different groups in Cambodia. Simple statistics indicate a significant temporary employment's wage penalty in the former, but show no noticeable difference in terms of wage rates between temporary and permanent Cambodian workers. These first observations suggest the particularity of temporary employment in Cambodia, which deserves further in-depth investigation.

From a gender perspective, except in Pakistan, males were more likely to engage in temporary jobs relative to females. In Cambodia, the incidences of temporary employment for men and women are roughly equal. Temporary employment was most popular among the youngest (Vietnam, the Philippines and Pakistan) and the oldest workers (Indonesia and Cambodia, to a lesser extent). It constituted up to 86 percent and 92 percent among youth in Vietnam and Pakistan, respectively. In the Philippines, the proportion of young workers engaging in temporary jobs was twice as much as for prime-aged workers, reaching 40 percent by 2009. By contrast, in Indonesia, older workers were twice more engaged in temporary employment than the youth (45 percent versus 22 percent).

In all cases except Cambodia, higher level of education was associated with lower probability of involving in temporary work. Vietnam serves as a typical example. Whereas only one quarter of highly qualified employees were working temporarily, almost all workers with no education were temporary (96 percent). Likewise, in Indonesia, workers with tertiary education and above virtually never took part in casual jobs (0.5 percent). In contrast, the most educated Cambodian employees were also the most engaging in temporary work, at 63 percent.

Public sector was less likely to recruit temporary workers than private one in Vietnam, the Philippines and Pakistan, while the reverse was true in Cambodia. In Pakistan, for instance, whilst almost all wage workers in the private sector were temporarily employed (93 percent), temporary workers accounted for only 12 percent of wage employees in the public sector.

In all countries except Cambodia, the incidence of temporary employment was the highest in agriculture and lowest in services. In Indonesia, for example, 71 percent of agricultural salaried jobs were temporary. Meanwhile, only seven percent of service wage workers engaged in a temporary job. Manufacturing is the second employer of temporary labor in Indonesia, at 24 percent. In Vietnam, the Philippines and Pakistan, temporary jobs were the second most prevalent in the trade sector, followed by the manufacturing sector. Meanwhile, the presence of temporary employment was more or less the same across different industries in Cambodia, though appearing slightly more prevalent in the service sector (57 percent).

Not surprisingly, lower-end occupations such as 'elementary occupations,' 'plant and machine operators and assemblers,' 'craft and related trade workers' tended to work temporarily more than superior occupations such as 'legislators, senior officials and managers' and 'professionals' in all studied countries excluding Cambodia. In the Philippines, 'elementary occupations' was the number-one temporary job provider, followed by 'craft and related trade workers.' The reverse was true for Pakistan. Meanwhile, 'skilled agricultural and fishery workers' in Indonesia were most inclined to temporary employment among all occupations. As expected, there was no remarkable differences in the incidence of temporary jobs among various occupations in Cambodia.

In consistence with the literature, wage workers in the informal economy were widely working temporarily. Notably, in Vietnam and Pakistan, almost all of informal sector wage jobs were also temporary (99 percent and 98 percent, respectively). This figure shows a close interrelation between informality and impermanence. The question of wage discrimination associated with temporary employment is therefore relevant to the dominant informal sector in these economies.

Table 2 also provides preliminary results on the temporary-permanent workers' wage differentials. In consistence with the findings in the empirical literature, the results show general remuneration disadvantages of temporary employment compared to permanent jobs. The raw wage gap appeared the most severe in Pakistan, where temporary workers earned only 42 percent as much as permanent ones did. In Indonesia, on average, temporary hourly wage was about half of the permanent one. Yet, the penalty appeared trivial in Cambodia, with a wage ratio of up to 98 percent.

In summary, temporary employment emerged as a noteworthy important phenomenon in a number of Asian countries, as a large proportion of wage workers derived all or part of their earnings from it. Descriptive statistics have revealed contrasting results in line with the aforementioned theoretical controversy: while in most of the cases, the disadvantaged in the labor market, such as youth, the less educated, those doing 'inferior' jobs, agricultural workers, workers in the private sector, were more likely to work temporarily; there was also evidence that temporary employment was not less favorable than permanent one. Given their prevalence across Asia, their diverse characteristics suggested by the statistical description, it is of great importance to finely evaluate the wage differentials associated with temporary employment relationships. In this attempt, the next section will present the empirical results derived from wage equations.

Table 1. Definition of Temporary Employment in National Labor Force Surveys

	Cambodia (2012)	Indonesia (2007)	Pakistan (2008-09)	Philippines (2009)	Vietnam (2007)
	Is the contract or agreement of...?	What is the status of (NAME) in his/her main job in the last one week?	What was the status of job's written contract/ agreement between the employee and the employer?	Nature of employment	Type of contract
Survey question to define temporary employment	1= Limited duration 2= Unlimited duration 3= Unspecified duration 4= Don't know	1. Self employed 2. Self-employed assisted by temporary/unpaid worker 3. Employer with permanent/paid workers 4. Employee 5. Casual employee in agriculture 6. Casual employee in non-agriculture 7. Unpaid worker	1. Permanent/ pensionable Job With contract/ agreement 2. Less than 1 year 3. Up to 3 years 4. Up to 5 years 5. Up to 10 years 6. 10 Years and more 7. Without contract/agreement	1-Permanent job/business/unpaid family work 2-Short-term or seasonal or casual job/business/unpaid family work 3-Worked for different employer on day to day or week to week basis	1- Indefinite-term labor contract 2- Definite-term labor contract 3- Verbal agreement 4- No contract
Temporary	1, 3, 4	5, 6	2, 3, 4, 5, 6, 7	2, 3	2, 3, 4

Source: Cambodia's Labor Force and Child Labor Survey 2012, Indonesia's LFS 2007-2008, Pakistan's LFS 2008-2009, the Philippines' LFS 2009, Vietnam's LFS 2007.

Table 2. Temporary Employees in Selected Asian Countries: Statistical Description

	Cambodia	Indonesia	Pakistan	Philippines	Vietnam
Wage workers, as % of total employment	46.03	38.45	36.83	53.19	30.86
Temporary workers, as % of wage workers	53.22	27.01	74.6	24.25	67.34
Temporary workers, as % of wage workers, by sex					
Male	53.82	28.35	74.14	25.54	71.39
Female	52.36	24.07	77.8	22.12	61.37
Temporary workers, as % of wage workers, by age					
15-24	52.67	22.1	91.6	39.63	85.64
25-54	53.2	26.32	67	19.63	62.01
55-64	57.2	45.44	69.37	16.39	60.18
Temporary workers, as % of wage workers, by level of education					
No education	52.6	71.87	91.94	39.3	95.84
Elementary	49.3	47.66	86.84	32.49	86.24
High school and vocational	59.35	14.8	57.28	28.37	55.28
University and more	63.08	0.47	30.67	14.22	25.69
Temporary workers, as % of wage workers, by sector					
Public	56.87	N.A.	14.52	10.34	28
Private	52.52	N.A.	93.01	26.76	89
Temporary workers, as % of wage workers, by industry					
Agriculture	52.59	71.37	97.93	44.57	89.86
Manufacturing	51.11	24.25	89.71	21.78	79.92
Trade	50.18	8.15	93.91	37.68	81.82
Services	56.47	7.35	48.16	17.86	41.13
Temporary workers, as % of wage workers, by occupational groups					
Legislators, senior officials and managers	53.87	0.06	42.16	4.58	N.A.
Professionals	63.42	0.52	48.67	6.64	N.A.
Technicians and associate professionals	49.9	1.69	38.08	15.44	N.A.

Clerks	62.67	0.45	25.16	15.27	N.A.
Service workers and shop and market sales workers	54.48	7.14	75.12	23.64	N.A.
Skilled agricultural and fishery workers	52.01	65.19	62.83	22.52	N.A.
Craft and related trades workers	47.51	29.46	92.49	25.48	N.A.
Plant and machine operators and assemblers	53.31	9.23	83.14	17.67	N.A.
Elementary occupations	51.16	48.49	87.82	35.85	N.A.
Temporary workers, as % of informal wage workers	49.02	N.A.	N.A.	N.A.	95.87
Temporary workers, as % of informal sector wage workers	46.57	N.A.	97.69	N.A.	99.28
Temporary/permanent wage ratio (hourly, in %)	98.09	49.12	42.3	64.73	60

Source: Cambodia's Labor Force and Child Labor Survey 2012, Indonesia's LFS 2007-2008, Pakistan's LFS 2008-2009, the Philippines' LFS 2009, Vietnam's LFS 2007. Authors' calculation.

3.3. Estimation strategies

This section aims to refine and to assess if temporary jobs are suffering or benefiting in financial terms after controlling for various personal, job-related and geographical characteristics. As in the standard Mincerian earnings equation, the general form of the wage equation for all workers is setup as:

$$y = Z\alpha + X\beta + \epsilon$$

where y denotes the logarithm of the hourly wage, Z is the dummy variable representing temporary employment, α accordingly captures the temporary employment's wage gap to be estimated. X is a set of explanatory variables considered as determining the wage rates, and β are their associated parameters. Finally, ϵ is the disturbance term or error variable.

X systematically includes four groups of variables as follows:

i. *Individual characteristics*: gender, head of household, marital status (all countries),[78] migration (Pakistan and Cambodia);[79] ethnicity (Vietnam).

ii. *Human capital*: age and its square, education level (all countries),[80] training (Pakistan),[81] and occupations (Cambodia, Indonesia, the Philippines and Pakistan).

As underlined in the literature (For examples, Edin and Richardson (2002); Albrecht et al. (2003)), the existence of wage gap is mostly caused by differences in human capital such as schooling, work experience, or tenure.

iii. *Job characteristics*: industries (all countries), public sector (Vietnam, the Philippines, Pakistan and Cambodia), informal sector (Vietnam, Pakistan and Cambodia), union member (Indonesia), business size (Vietnam).

According to Arulampalam et al. (2007), it is important to control for job characteristics as they might be endogenous. In addition, job controls could also reflect unmeasured human capital.

iv. *Geographical characteristics*: rural area, provinces/regions (all countries).

To assure the cross-country comparability, in the first step, the benchmark model includes only mutual variables that exist in all national LFS of selected countries. Extended models including additional country-specific information are then implemented for more information.

To account for the self-selection into wage employment as well as to assess the stability of the estimation results across different econometric models, Heckman procedure is applied. By definition, the wage rate can only be observed for wage workers, who might not be representative

[78] Single is selected as the reference group.

[79] No migrant is selected as the reference group.

[80] No education is taken as the reference group.

[81] No training is used as the reference group.

of the whole labor force. The Heckman selection model could correct for this bias if there are some identifying variables which strongly affect one's probability of being a wage worker but not their wage level. Practical examples include household variables such as income of the spouse, household wealth, non-labor household income, children (Puhani, 2000), whether parents are living together or are living in the same neighborhood (Nawata, 2004) and marital status.[82] Given the available dataset, marital status and whether the individual is head of their household are chosen as identifying variables in our model. Married household heads, for example, possibly bear more pressure to work than the others, but the fact that they are married and heads of household hardly has a significant impact on their offered wage.

Although the Heckman selection model is widely applied in economics, one should keep in mind its drawbacks. First, it is quite sensitive to the specification of the model, much more so than ordinary estimation.[83] Second, excluded variables are difficult to find, and in reality there is no such variables since anything that determines the working decision would also affect the wages through the labor supply (Fathollah & Orhan, 2008). Without such identifying variables, the model is only identified through the non-linearity of the inverse Mills ratio lambda, which leads to non-robust results in most of the cases due to collinearity problems (Puhani, 2000). If the collinearity problem cannot be solved in empirical studies, standard OLS is recommended (ibid). Moreover, although Heckman procedure is sometimes considered 'indispensable' whenever sample representativeness is questionable, Heckman himself acknowledged that his estimator is only to "provide good starting values for maximum likelihood estimation" and "exploratory empirical work" (Heckman, 1979). In fact, evidence suggests that sample selection bias is typically modest (Hyclak et al., 2005). Therefore, the results obtained from the Heckit method in this paper are primarily used to examine the robustness of OLS estimates.

Besides measuring wage gaps at mean wage level, estimation of quantile regression is also provided in order to allow the wage gaps between temporary and permanent jobs to differ along the wage distribution. Quantile wage regressions consider specific parts of the conditional distribution of the hourly wage and indicate the influence of the different explanatory variables on conditional wage respectively at the bottom, at the median and at the top of the distribution. Using the aforementioned notation, the general quantile regression model can be written as:

$$q_{\varrho}(y) = Z\alpha(\varrho) + X\beta(\varrho) + \epsilon, \varrho \in [0,1]$$

where $q_{\varrho}(y)$ is the ϱ^{th} conditional logarithm of hourly wage quantile. The set of coefficients $\beta(\varrho)$ provides the estimated rates of return to the different covariates at the ϱ^{th} quantile of the log wage distribution and the coefficients $\alpha(\varrho)$ measure the parts of the wage differentials that are due to job-form difference at the various quantiles. In a quantile regression, the distribution of the error term is left unspecified.

Using quantile regression method provides robust estimates, particularly for misspecification errors related to non-normality and heteroscedasticity. Another important

[82] Source: stata.com

[83] Source: stata.com

advantage of quantile regression applied in wage gap analysis is the identification of distributional effects, permitting to test two opposite hypotheses, namely, *sticky floor* and *glass ceiling*. *Sticky floor* effect can be measured by three indicators: (i) 10-all gaps, (ii) 10-25 difference, or (iii) 10-50 difference. Likewise, three criteria can be used to determine the existence of *glass ceiling* phenomenon: (i) 90-all gaps, (ii) 90-75 difference, or (iii) 90-50 difference.

4. Econometric analysis of wage differentials

4.1. Estimates at mean: OLS and Heckman models

Table 3 presents the estimation results of benchmark OLS model with a similar set of covariates in all selected countries. Diverse results appear, in consistence with controversial predictions in the aforementioned theories. On the one hand, in Indonesia, Pakistan, the Philippines and Vietnam, temporary workers face a wage penalty compared to their permanent counterparts. The temporary workers' wage penalty is the most severe in Pakistan (45 percent) and Vietnam (31 percent), and fairly modest in Indonesia (16 percent) and the Philippines (seven percent). On the other hand, Cambodia shows a temporary workers' wage premium, at six percent.[84]

In general, the included individual characteristics, human capital, job and geographical variables are significantly correlated with workers' wage. In all countries, females seem to earn less than males, *ceteris paribus*, and the gender wage gap is the most pronounced in Pakistan (33 percent). Being head of a household could have a small positive relation to wage rate in Indonesia and Vietnam (nearly four and two percent, respectively). In general, married workers earn a higher wage than the single; while in some cases, widowed and divorced workers are probably remunerated less. Experience and seniority, approximated by age, appear to be positively correlated with individual wage. One more year of age is associated with an around three to four percent increase in the wage rate, other things equal. In line with the Human Capital theory, obtaining some kind of education/qualification instead of nothing could raise individual wage. Moreover, the higher the education level, the higher wage. In the Philippines, for example, while elementary schooling is associated with an 18 percent increase in individual wages, university-and-above graduates could earn twice as much as the uneducated do, *ceteris paribus*. In all selected countries, manufacturing wage jobs are the best remunerated among all industries. On average, rural workers are worse paid than urban counterparts in Cambodia, Pakistan, the Philippines and Vietnam. The opposite is observed in Indonesia, but the wage difference between two groups is modest (two percent).

[84] For the wage gaps to be meaningful, wages need to distribute along a sufficiently wide range. The results show large deviations around the mean wages of wage workers in all five selected countries. The means and standard deviations of wage are: Cambodia: 2439.58 & 4577.87; Indonesia: 5818.03 & 7795.68; Pakistan: 46.33 & 81.49; the Philippines: 37.57 & 40.55; and Vietnam: 7.69 & 5.73.

Table 3. Temporary Workers' Wage Gap in Selected Asian Countries – Benchmark OLS Model

VARIABLES	(1) **Cambodia**	(2) **Indonesia**	(3) **Pakistan**	(4) **Philippines**	(5) **Vietnam**
Temporary Employee	0.056***	-0.158***	-0.446***	-0.073***	-0.309***
	(0.015)	(0.004)	(0.012)	(0.007)	(0.004)
Individual characteristics					
Gender Female	-0.122***	-0.197***	-0.327***	-0.224***	-0.144***
	(0.015)	(0.004)	(0.016)	(0.008)	(0.003)
Head of household	0.004	0.039***	0.001	0.007	0.019***
	(0.025)	(0.004)	(0.012)	(0.009)	(0.004)
Marital status (reference: single)					
Married	0.023	0.101***	0.044***	0.107***	0.021***
	(0.022)	(0.005)	(0.013)	(0.009)	(0.004)
Widowed	-0.060	-0.096***	-0.090**	-0.015	-0.058***
	(0.055)	(0.010)	(0.038)	(0.023)	(0.013)
Divorced	-0.101*	-0.080***	-0.016	-0.021	-0.008
	(0.058)	(0.010)	(0.066)	(0.025)	(0.011)
Human capital					
Age	0.034***	0.039***	0.030***	0.029***	0.042***
	(0.004)	(0.001)	(0.003)	(0.002)	(0.001)
Age squared	-0.000***	-0.000***	-0.000***	-0.000***	-0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Education (reference: no education)					
Primary and secondary	0.062***	0.169***	0.090***	0.184***	0.061***
	(0.015)	(0.009)	(0.010)	(0.039)	(0.006)
High school	0.247***	0.492***	0.284***	0.359***	0.226***
	(0.031)	(0.009)	(0.011)	(0.039)	(0.007)
University and above	0.583***	1.163***	0.831***	1.070***	0.576***
	(0.058)	(0.010)	(0.017)	(0.039)	(0.008)
Job characteristics					
Industry (reference: agriculture)					
Manufacturing	0.066***	0.151***	0.210***	0.123***	0.050***
	(0.018)	(0.004)	(0.016)	(0.012)	(0.006)
Trade	0.080	-0.038***	-0.089***	0.224***	0.036***
	(0.053)	(0.006)	(0.021)	(0.010)	(0.008)
Services	0.017	0.200***	0.109***	0.011	0.023***
	(0.023)	(0.005)	(0.017)	(0.010)	(0.007)
Geographical characteristics					
Rural area	-0.132***	0.023***	-0.063***	-0.067***	-0.062***
	(0.031)	(0.003)	(0.008)	(0.008)	(0.004)
Region	Yes	Yes	Yes	Yes	Yes
Constant	7.172***	7.029***	2.829***	2.081***	0.936***
	(0.078)	(0.017)	(0.048)	(0.050)	(0.020)
Observations	9,050	154,074	25,460	33,530	117,968
R-squared	0.099	0.497	0.393	0.435	0.366

*Source: Cambodia's Labor Force and Child Labor Survey 2012, Indonesia's LFS 2007-2008, Pakistan's LFS 2008-2009, the Philippines' LFS 2009, Vietnam's LFS 2007. Authors' calculation. Robust standard errors in parentheses; ***p<0.01, **p<0.05, *p<0.1.*

The first columns of **Table 6**, **Table 7**, **Table 8**, **Table 9** and **Table 10** describe OLS estimates of temporary-permanent wage differentials in extended models which further include country-specific explanatory variables. In Cambodia, the extended model introduces additional controls for migration, occupation, institutional sector, business size and trade union membership. Compared to the benchmark model, the estimated wage gap slightly decreases (from six percent to five percent). There also appears a relationship between migration status and individual wage. In comparison to the non-migrants, migrated workers from other provinces tend to achieve a higher wage level (eleven percent). Notably, the coefficient associated with manufacturing sector turns into non-significant, while trade coefficient becomes positively significant. Public sector jobs seem to offer remarkably lower wages than comparable jobs in the private sector (minus 21 percent). On average, there is no significant relationship between trade union membership and workers' wage.[85]

In Indonesia, by including occupation, trade union membership and social security in the regression model, the estimated wage penalty drops threefold, from 16 percent to five percent. Members of trade unions tend to obtain a higher wage than non-members by ten percent, *ceteris paribus*. Indonesian workers who are covered by social security, as an indicator for formal jobs, are much better remunerated than the uncovered counterparts (34 percent).

In Pakistan, the temporary worker's wage penalty is reduced by half (from 45 percent to 26 percent) when migration, training, occupation and institutional sectors are further controlled for. Compared to the non-migrants, on average, workers who have migrated less than one year receive a nine percent higher wage rate, and those who have migrated for at least ten years are better paid by six percent. This is perhaps because migration decision is principally determined by wages, mainly for those who have just moved and those who have settled down in a new place. Attending job training is associated with a higher wage rate, and on-the-job training appears to have a larger impact than off-the-job training. In contrast to Cambodia, public sector jobs in Pakistan are generally better remunerated than private sector employment by 28 percent. Other things equal, jobs in the informal sector bring a slightly lower wage than those in the formal sector.

In the Philippines, by controlling for occupational categories and institutional sector, the wage differential between temporary and permanent workers is greatly attenuated from seven percent to two percent. Notably, the impact of tertiary education on individual wage is reduced by half, possibly due to the inclusion of occupation and working sectors. In addition, manufacturing modality turns into non-significant while service sector jobs become significantly less remunerated than farming work. On average, workers in the public sector earn a wage rate of 25 percent higher than in the private sector, *ceteris paribus*.

By contrast, there is no noticeable change in the estimated wage gap in Vietnam when additional demographic and job characteristics are included in the model (from 31 percent to 29 percent). Foreign sector, interestingly, provides lower paid jobs than the domestic private sector on average.

A summary and comparison of benchmark and extended models are also provided in **Table 4**. The different changes between the benchmark OLS model and the extended OLS model

[85] When detailed quantiles are considered, significant estimates appear.

in different countries could be attributed to different sets of additional variables from country to country. The stable results in Vietnam are possibly related to the fact that there is no available information on occupation in Vietnam, which serves as an important explaining factor of individual wages and could be highly correlated with employment status as in other countries' extended models. By contrast, in Pakistan, where the difference in the estimated wage gap between the two models is the largest, both occupation and training are controlled for. Whether a worker obtains a training, whether this is an on-the-job training or off-the-job training is very likely to be strongly related to their employment status, for that the inclusion of training may further exaggerate the difference between the two estimates. Likewise, in Indonesia, the presence of trade union membership in the extended model could be partially responsible for the drastic fall in the estimated wage differential.

The question of self-selection into wage employment as well as the reliability of OLS estimates are addressed with the Heckman procedure as presented in **Table 5**, using a same set of covariates for all studied countries. 'Stage 2' means the main equation and 'stage 1' indicates the selection equation. In general, the 'corrected' estimates almost stay the same as what obtained from the benchmark OLS models, and all control variables keep the same signs and close magnitudes despite the presence of sample selection in Indonesia, Pakistan, the Philippines and Vietnam. In Cambodia, in fact, no such sample selection exists (lambda is not significant). This highlights that controlling for marital status and household head in OLS regression yields not much different results from applying the Heckit method with marital status and household head serving as identifying variables in the first stage. Therefore, the next section does not apply Heckman correction method in the Quantile Regression framework.

4.2. Quantile estimation analysis

Table 4 summarizes the quantile regression estimates from the benchmark and extended models in five selected countries. **Table 6**, **Table 7**, **Table 8**, **Table 9** and **Table 10** provide detailed quantile estimation results using the extended set of control variables in each country. The variation of temporary-permanent workers' wage differentials along the wage distribution is illustrated in **Figure 1**. Interestingly, both benchmark and extended models show that there are diverse distribution patterns of temporary-permanent workers' wage gap in the selected Asian countries. In Vietnam and Pakistan, the wage penalty is greater in the lower tiers of the wage distribution, suggesting a *sticky floor* effect. Meanwhile, a *glass ceiling* effect seems to be relevant in case of Indonesia. The temporary workers' wage penalty in the Philippines and the temporary worker's wage premium in Cambodia are fairly stable across different percentiles of the wage profile.

In Vietnam, the wage gap continuously decreases from 32 percent in the 10^{th} quantile to 19 percent in the 90^{th} quantile. Pakistan experiences the same tendency, with the gap smoothly decreasing from 31 percent at the bottom to 14 percent at the top of the wage distribution. The results for Vietnam and Pakistan support the *sticky floor* hypothesis, and the 10-all gap criterion is relevant in both countries. This suggests that the penalty of being in temporary jobs could be more severe for disadvantaged workers in Vietnam and Pakistan, who are in the lower tiers of the wage distribution, than for high wage earners.

By contrast, the temporary workers' wage penalty in Indonesia gradually increases from three percent (10th percentile) to seven percent (80th percentile), before falling to five percent in the remaining percentiles of the wage distribution. The hypothesis of *glass ceiling* is therefore relevant in case of Indonesia. However, three indicators mentioned in Arulampalam et al. (2007) and Christofides et al. (2013) (i.e. 90-all gaps, 90-75 difference and 90-50 difference) are no longer satisfied in Indonesia. The *glass ceiling* phenomenon in this country seems to be associated with the '75-all gap,' that is the wage penalty wider at the 75th and narrower at both the bottom and the top of the conditional wage distribution.

In the Philippines and Cambodia, no clear distributional effects of the penalty/premium associated with temporary employment status can be found along the pay ladder. Although the point estimates seem to exhibit an upward trend graphically, the confidence intervals are so large that no significant distributional effect could be assured. This indicates, in other words, that temporary workers in the Philippines and Cambodia suffer/benefit financially similar levels of penalty/premium at any positions in the wage distribution.

Table 4. Temporary-Permanent Workers' Wage Gap: OLS and Quantile Regression

Variables	Benchmark model						Extended model					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	OLS	QR .10	QR .25	QR .50	QR .75	QR .90	OLS	QR .10	QR .25	QR .50	QR .75	QR .90
Cambodia	0.056***	0.030**	0.024**	0.035***	0.046***	0.059***	0.047***	0.025**	0.036***	0.028***	0.046***	0.054***
	(0.015)	(0.014)	(0.010)	(0.009)	-0.011	(0.019)	(0.015)	(0.011)	(0.010)	(0.009)	(0.010)	(0.019)
Indonesia	-0.158***	-0.119***	-0.138***	-0.169***	-0.192***	-0.180***	-0.047***	-0.029***	-0.036***	-0.047***	-0.070***	-0.068***
	(0.004)	(0.007)	(0.005)	(0.005)	(0.005)	(0.006)	(0.004)	(0.007)	(0.005)	(0.004)	(0.005)	(0.006)
Pakistan	-0.446***	-0.527***	-0.511***	-0.464***	-0.388***	-0.268***	-0.256***	-0.314***	-0.300***	-0.265***	-0.204***	-0.140***
	(0.012)	(0.017)	(0.012)	(0.009)	(0.010)	(0.013)	(0.016)	(0.016)	(0.015)	(0.010)	(0.012)	(0.015)
Philippines	-0.073***	-0.049***	-0.066***	-0.081***	-0.088***	-0.060***	-0.016**	-0.042***	-0.035***	-0.024***	-0.027***	-0.022***
	(0.007)	(0.011)	(0.009)	(0.007)	(0.008)	(0.008)	(0.007)	(0.010)	(0.008)	(0.006)	(0.006)	(0.008)
Vietnam	-0.309***	-0.329***	-0.311***	-0.285***	-0.265***	-0.233***	-0.289***	-0.320***	-0.293***	-0.251***	-0.221***	-0.185***
	(0.004)	(0.007)	(0.005)	(0.004)	(0.004)	(0.006)	(0.005)	(0.007)	(0.005)	(0.004)	(0.005)	(0.007)

Source: Cambodia's Labor Force and Child Labor Survey 2012, Indonesia's LFS 2007-2008, Pakistan's LFS 2008-2009, the Philippines' LFS 2009, Vietnam's LFS 2007. Authors' calculation.
*Robust standard errors in parentheses; ***p<0.01, **p<0.05, *p<0.1.*

Table 5. Temporary-Permanent Workers' Wage Gap: Heckman Correction for Sample Selection

	Cambodia		Indonesia		Pakistan		Philippines		Vietnam	
	(1)	(2)	(9)	(10)	(7)	(8)	(5)	(6)	(3)	(4)
	Stage 2	Stage 1	Stage 2	Stage 1	Stage 2	Stage 1	Stage 2	Stage 1	Stage 2	Stage 1
Temporary employment	0.055***		-0.160***		-0.448***		-0.076***		-0.312***	
	(0.015)		(0.004)		(0.011)		(0.008)		(0.004)	
Individual characteristics										
Female	-0.139***	-0.190***	-0.250***	-0.139***	-0.391***	-0.246***	-0.231***	0.087***	-0.114***	-0.220***
	(0.016)	(0.026)	(0.004)	(0.006)	(0.022)	(0.019)	(0.007)	(0.013)	(0.004)	(0.006)
Head of household				0.130***		0.066***		0.093***		-0.008
				(0.007)		(0.017)		(0.014)		(0.007)
Marital status (reference: single)										
Married		-0.490***		-0.301***		-0.137***		-0.276***		-0.190***
		(0.034)		(0.008)		(0.020)		(0.015)		(0.009)
Widowed		-0.133*		-0.213***		0.037		-0.213***		-0.140***
		(0.074)		(0.018)		(0.053)		(0.030)		(0.020)
Divorced		-0.310***		-0.274***		0.062		-0.148***		-0.082***
		(0.087)		(0.015)		(0.101)		(0.039)		(0.022)
Human capital										
Age	0.033***	-0.031***	0.051***	0.010***	0.040***	0.035***	0.037***	0.004	0.047***	0.007***
	(0.004)	(0.005)	(0.001)	(0.001)	(0.003)	(0.004)	(0.001)	(0.002)	(0.001)	(0.001)
Age squared	-0.000***	0.000	-0.000***	-0.000***	-0.000***	-0.001***	-0.000***	-0.000***	-0.000***	-0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Education (reference: no education)										
Primary and secondary	0.062***	-0.038	0.195***	0.045***	0.061***	-0.134***	0.211***	0.121***	0.092***	-0.158***
	(0.016)	(0.023)	(0.010)	(0.012)	(0.014)	(0.015)	(0.039)	(0.044)	(0.007)	(0.010)
High school	0.260***	0.266***	0.537***	0.234***	0.276***	-0.037**	0.388***	0.100**	0.160***	0.368***
	(0.030)	(0.047)	(0.011)	(0.012)	(0.012)	(0.017)	(0.039)	(0.044)	(0.008)	(0.011)
University and above	0.619***	0.979***	1.285***	1.418***	0.879***	0.282***	1.136***	0.506***	0.359***	1.532***

	(0.055)	(0.106)	(0.017)	(0.015)	(0.022)	(0.024)	(0.041)	(0.045)	(0.015)	(0.014)
Job characteristics										
Industry (reference: agriculture)										
Manufacturing	0.122**	1.172***	0.255***	1.100***	0.761***	2.020***	0.195***	0.542***	-0.310***	1.642***
	(0.050)	(0.028)	(0.015)	(0.006)	(0.173)	(0.018)	(0.020)	(0.017)	(0.021)	(0.008)
Trade	0.044	-0.646***	-0.055***	-0.242***	0.084	0.523***	0.416***	2.131***	-0.037***	0.173***
	(0.048)	(0.040)	(0.007)	(0.008)	(0.059)	(0.020)	(0.046)	(0.032)	(0.008)	(0.010)
Services	0.071	1.090***	0.281***	0.807***	0.658***	2.010***	0.081***	0.542***	-0.275***	1.267***
	(0.049)	(0.031)	(0.012)	(0.006)	(0.173)	(0.018)	(0.020)	(0.014)	(0.018)	(0.008)
Geographic characteristics										
Rural area	-0.124***	0.186***	0.009**	-0.226***	-0.063***	-0.007	-0.086***	-0.193***	-0.050***	-0.057***
	(0.028)	(0.042)	(0.004)	(0.005)	(0.009)	(0.014)	(0.009)	(0.012)	(0.004)	(0.006)
Region	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Constant	7.125***	0.700***	6.682***	-0.588***	1.990***	-1.741***	1.765***	-0.330***	1.363***	-0.966***
	(0.072)	(0.111)	(0.027)	(0.025)	(0.245)	(0.058)	(0.063)	(0.063)	(0.032)	(0.028)
Lambda	0.086	0.086	0.137***		0.388***	0.388***	0.173***	0.173***	-0.318***	-0.318***
	(0.069)	(0.069)	(0.018)		(0.121)	(0.121)	(0.040)	(0.040)	(0.018)	(0.018)
Observations	18,546		485,785		69,493		78,932		334,445	

Source: Cambodia's Labor Force and Child Labor Survey 2012, Indonesia's LFS 2007-2008, Pakistan's LFS 2008-2009, the Philippines' LFS 2009, Vietnam's LFS 2007. Authors' calculation.
*Robust standard errors in parentheses; ***p<0.01, **p<0.05, *p<0.1.*

Table 6. Temporary-Permanent Workers' Wage Gap: Extended OLS and Quantile Regression - Cambodia.

VARIABLES	(1) OLS	(2) QR .10	(3) QR .25	(4) QR .50	(5) QR .75	(6) QR .90
Temporary Employee	0.047***	0.025**	0.036***	0.028***	0.046***	0.054***
	(0.015)	(0.011)	(0.010)	(0.009)	(0.010)	(0.019)
Individual characteristics						
Gender Female	-0.146***	-0.111***	-0.086***	-0.104***	-0.103***	-0.188***
	(0.016)	(0.015)	(0.013)	(0.010)	(0.011)	(0.022)
Migration (reference: non migrant)						
Migrant from another village in this province	-0.003	0.010	0.037**	0.018	0.024	0.046
	(0.024)	(0.023)	(0.018)	(0.015)	(0.016)	(0.032)
Migrant from another province	0.105***	0.023	0.066***	0.074***	0.102***	0.207***
	(0.026)	(0.026)	(0.023)	(0.014)	(0.024)	(0.039)
Migrant from another country	0.109	0.180**	0.111	0.223	0.121	-0.020
	(0.140)	(0.079)	(0.131)	(0.254)	(0.084)	(0.079)
Marital status (reference: single)						
Married	0.022	-0.003	0.018	0.025*	0.037***	0.069**
	(0.022)	(0.017)	(0.016)	(0.013)	(0.014)	(0.029)
Widowed	-0.047	-0.123	-0.012	-0.034	-0.016	-0.003
	(0.054)	(0.103)	(0.058)	(0.028)	(0.033)	(0.049)
Divorced	-0.100*	-0.047**	-0.070	-0.054**	-0.020	0.008
	(0.056)	(0.022)	(0.057)	(0.026)	(0.024)	(0.040)
Head of household	0.017	0.075***	0.026	0.017	0.023	-0.016
	(0.025)	(0.023)	(0.018)	(0.014)	(0.018)	(0.034)
Human capital						
Age	0.027***	0.025***	0.021***	0.016***	0.012***	0.008
	(0.004)	(0.004)	(0.003)	(0.003)	(0.003)	(0.005)
Age squared	-0.000***	-0.000***	-0.000***	-0.000***	-0.000***	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Education (reference: no education)						
Primary and secondary	0.052***	0.082***	0.067***	0.036***	0.027***	0.031
	(0.015)	(0.012)	(0.011)	(0.009)	(0.010)	(0.020)
High school	0.199***	0.218***	0.200***	0.134***	0.127***	0.146***

		(0.033)	(0.043)	(0.023)	(0.017)	(0.032)	(0.049)
	University and above	0.528***	0.528***	0.490***	0.553***	0.620***	0.459***
		(0.061)	(0.074)	(0.069)	(0.044)	(0.067)	(0.064)
Occupation		Yes	Yes	Yes	Yes	Yes	Yes
Job characteristics							
Industry (reference: agriculture)							
	Manufacturing	0.022	0.089***	0.065***	0.066***	0.029**	-0.022
		(0.022)	(0.023)	(0.018)	(0.013)	(0.014)	(0.033)
	Trade	0.126**	0.033	-0.025	0.065	0.149***	0.311***
		(0.055)	(0.060)	(0.052)	(0.043)	(0.054)	(0.114)
	Services	0.028	-0.086***	-0.059**	0.047*	0.160***	0.184***
		(0.029)	(0.025)	(0.027)	(0.024)	(0.025)	(0.048)
Public sector		-0.209***	-0.119***	-0.113***	-0.162***	-0.173***	-0.213***
		(0.039)	(0.031)	(0.031)	(0.025)	(0.030)	(0.045)
Business size		Yes	Yes	Yes	Yes	Yes	Yes
Trade union membership		-0.019	0.039***	0.018	0.017	-0.018	-0.085***
		(0.020)	(0.013)	(0.013)	(0.014)	(0.014)	(0.028)
Geographical characteristics							
Rural area		-0.120***	-0.052***	-0.080***	-0.079***	-0.051**	-0.107***
		(0.031)	(0.020)	(0.020)	(0.016)	(0.024)	(0.040)
Region		Yes	Yes	Yes	Yes	Yes	Yes
Constant		7.349***	6.647***	6.996***	7.530***	7.732***	8.259***
		(0.102)	(0.095)	(0.097)	(0.074)	(0.082)	(0.165)
Observations		9,050	9,050	9,050	9,050	9,050	9,050
R-squared		0.129					

*Source: Cambodia's Labor Force and Child Labor Survey 2012. Authors' calculation. Robust standard errors in parentheses; ***p<0.01, **p<0.05, *p<0.1.*

Table 7. Temporary-Permanent Workers' Wage Gap: Extended OLS and Quantile Regression – Indonesia

VARIABLES	(1) OLS	(2) QR .10	(3) QR .25	(4) QR .50	(5) QR .75	(6) QR .90
Temporary Employee	-0.047***	-0.029***	-0.036***	-0.047***	-0.070***	-0.068***
	(0.004)	(0.007)	(0.005)	(0.004)	(0.005)	(0.006)
Individual characteristics						
Gender Female	-0.227***	-0.268***	-0.258***	-0.219***	-0.189***	-0.181***
	(0.004)	(0.006)	(0.005)	(0.005)	(0.005)	(0.005)
Head of household	0.023***	0.028***	0.026***	0.026***	0.024***	0.021***
	(0.004)	(0.006)	(0.005)	(0.005)	(0.005)	(0.006)
Marital status (reference: single)						
Married	0.079***	0.083***	0.077***	0.071***	0.068***	0.076***
	(0.004)	(0.007)	(0.006)	(0.005)	(0.005)	(0.006)
Widowed	-0.047***	0.071***	-0.007	-0.051***	-0.102***	-0.164***
	(0.010)	(0.010)	(0.014)	(0.013)	(0.012)	(0.010)
Divorced	-0.040***	0.041***	-0.007	-0.056***	-0.086***	-0.116***
	(0.009)	(0.012)	(0.013)	(0.012)	(0.012)	(0.015)
Human capital						
Age	0.030***	0.031***	0.031***	0.028***	0.027***	0.028***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Age squared	-0.000***	-0.000***	-0.000***	-0.000***	-0.000***	-0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Education (reference: no education)						
Primary and secondary	0.139***	0.033***	0.114***	0.145***	0.162***	0.230***
	(0.008)	(0.012)	(0.012)	(0.010)	(0.010)	(0.009)
High school	0.317***	0.234***	0.292***	0.313***	0.320***	0.385***
	(0.009)	(0.012)	(0.012)	(0.011)	(0.011)	(0.010)
University and above	0.705***	0.682***	0.689***	0.664***	0.652***	0.753***

		(0.010)	(0.014)	(0.014)	(0.012)	(0.012)	(0.012)
Occupation		Yes	Yes	Yes	Yes	Yes	Yes
Job characteristics							
Industry (reference: agriculture)							
	Manufacturing	0.145***	0.172***	0.168***	0.163***	0.134***	0.098***
		(0.005)	(0.009)	(0.007)	(0.006)	(0.006)	(0.008)
	Trade	0.014*	0.043***	0.038***	0.031***	0.002	-0.039***
		(0.008)	(0.012)	(0.009)	(0.009)	(0.009)	(0.012)
	Services	0.062***	0.041***	0.042***	0.071***	0.086***	0.097***
		(0.006)	(0.009)	(0.007)	(0.007)	(0.007)	(0.009)
Trade union membership		0.099***	0.097***	0.095***	0.098***	0.089***	0.087***
		(0.005)	(0.008)	(0.006)	(0.005)	(0.005)	(0.007)
Formal employment (Social security)		0.343***	0.453***	0.412***	0.336***	0.274***	0.223***
		(0.004)	(0.006)	(0.005)	(0.004)	(0.004)	(0.005)
Geographical characteristics							
Rural area		0.027***	0.013**	0.022***	0.032***	0.036***	0.024***
		(0.003)	(0.005)	(0.004)	(0.004)	(0.004)	(0.004)
Provinces		Yes	Yes	Yes	Yes	Yes	Yes
Constant		7.873***	7.262***	7.572***	7.951***	8.239***	8.392***
		(0.020)	(0.038)	(0.027)	(0.024)	(0.023)	(0.025)
Observations		154,074	154,074	154,074	154,074	154,074	154,074
R-squared		0.568					

*Source: Indonesia's LFS 2007-2008. Authors' calculation. Robust standard errors in parentheses; ***p<0.01, **p<0.05, *p<0.1.*

Table 8. Temporary-Permanent Workers' Wage Gap: Extended OLS and Quantile Regression – Pakistan

VARIABLES	(1) OLS	(2) QR .10	(3) QR .25	(4) QR .50	(5) QR .75	(6) QR .90
Temporary Employee	-0.256***	-0.314***	-0.300***	-0.265***	-0.204***	-0.140***
	(0.016)	(0.016)	(0.015)	(0.010)	(0.012)	(0.015)
Individual characteristics						
Gender Female	-0.343***	-0.554***	-0.498***	-0.339***	-0.228***	-0.153***
	(0.017)	(0.027)	(0.020)	(0.017)	(0.016)	(0.019)
Head of household	-0.008	0.007	0.012	0.001	-0.005	-0.021*
	(0.011)	(0.012)	(0.012)	(0.009)	(0.010)	(0.011)
Marital status (reference: single)						
Married	0.039***	0.065	0.022	0.017	0.038***	0.019
	(0.013)	(0.000)	(0.013)	(0.011)	(0.012)	(0.015)
Widowed	-0.067*	-0.056	-0.043	-0.066*	-0.041	-0.094***
	(0.037)	(0.051)	(0.034)	(0.037)	(0.042)	(0.019)
Divorced	-0.013	0.011	0.075*	-0.030	-0.020	-0.144***
	(0.066)	(0.134)	(0.045)	(0.046)	(0.032)	(0.021)
Migration (reference: non migrant)						
Migrant less than 1 year	0.085**	0.001	0.052	0.091***	0.065	0.158**
	(0.043)	(0.041)	(0.050)	(0.026)	(0.068)	(0.064)
Migrant from 2-4 years	0.028	0.024	0.068**	0.040**	-0.001	0.007
	(0.027)	(0.042)	(0.030)	(0.019)	(0.025)	(0.045)
Migrant from 5- 9 years	0.032	0.030	0.032	0.040***	0.014	0.012
	(0.027)	(0.035)	(0.024)	(0.014)	(0.023)	(0.037)
Migrant since 10 years and over	0.064***	0.053**	0.068***	0.078***	0.050***	0.029*
	(0.014)	(0.024)	(0.017)	(0.011)	(0.010)	(0.015)
Human capital						
Age	0.030***	0.041	0.033***	0.027***	0.023***	0.021***
	(0.003)	(0.000)	(0.003)	(0.002)	(0.002)	(0.003)
Age squared	-0.000***	-0.000***	-0.000***	-0.000***	-0.000***	-0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Education (reference: no education)						
Primary and secondary	0.064***	0.075***	0.070***	0.075***	0.059***	0.065***
	(0.010)	(0.015)	(0.010)	(0.008)	(0.009)	(0.011)
High school	0.191***	0.180	0.162***	0.176***	0.166***	0.196***
	(0.013)	(0.000)	(0.013)	(0.009)	(0.010)	(0.015)
University and above	0.595***	0.452***	0.530***	0.577***	0.575***	0.635***

		(0.020)	(0.026)	(0.020)	(0.015)	(0.016)	(0.019)
Training (reference: no training)							
	On-the-job training	0.070***	0.093***	0.085***	0.059***	0.065***	0.041**
		(0.017)	(0.028)	(0.019)	(0.013)	(0.019)	(0.018)
	Off-the-job training	0.029*	0.001	0.014	0.010	0.052***	0.080***
		(0.017)	(0.021)	(0.021)	(0.016)	(0.017)	(0.018)
Occupation		Yes	Yes	Yes	Yes	Yes	Yes
Job characteristics							
Industry (reference: agriculture)							
	Manufacturing	0.190***	0.302	0.286***	0.250***	0.169***	0.081***
		(0.018)	(0.000)	(0.019)	(0.016)	(0.018)	(0.025)
	Trade	-0.068***	-0.013	-0.008	-0.008	-0.062***	-0.101***
		(0.024)	(0.030)	(0.027)	(0.021)	(0.023)	(0.029)
	Services	0.050**	0.093***	0.119***	0.114***	0.070***	0.021
		(0.020)	(0.013)	(0.021)	(0.017)	(0.019)	(0.027)
Public sector		0.275***	0.338	0.287***	0.243***	0.204***	0.157***
		(0.017)	(0.000)	(0.016)	(0.011)	(0.013)	(0.016)
Informal sector		-0.048***	-0.043***	-0.057***	-0.053***	-0.039***	-0.055***
		(0.010)	(0.015)	(0.012)	(0.009)	(0.010)	(0.012)
Geographical characteristics							
Rural area		-0.046***	-0.009	-0.036***	-0.032***	-0.052***	-0.054***
		(0.008)	(0.000)	(0.009)	(0.007)	(0.007)	(0.009)
Region		Yes	Yes	Yes	Yes	Yes	Yes
Constant		3.114***	1.899	2.683***	3.225***	3.773***	4.183***
		(0.054)	(0.000)	(0.057)	(0.050)	(0.046)	(0.064)
Observations		25,459	25,459	25,459	25,459	25,459	25,459
R-squared		0.419					

*Source: Pakistan's LFS 2008-2009. Authors' calculation. Robust standard errors in parentheses; ***p<0.01, **p<0.05, *p<0.1.*

Table 9. Temporary-Permanent Workers' Wage Gap: Extended OLS and Quantile Regression – The Philippines

VARIABLES	(1) OLS	(2) QR .10	(3) QR .25	(4) QR .50	(5) QR .75	(6) QR .90
Temporary Employee	-0.016**	-0.042***	-0.035***	-0.024***	-0.027***	-0.022***
	(0.007)	(0.010)	(0.008)	(0.006)	(0.006)	(0.008)
Individual characteristics						
Gender Female	-0.249***	-0.290***	-0.283***	-0.231***	-0.161***	-0.149***
	(0.007)	(0.012)	(0.009)	(0.008)	(0.007)	(0.009)
Head of household	0.017**	0.037***	0.019**	0.015**	0.010	0.020**
	(0.008)	(0.012)	(0.009)	(0.007)	(0.006)	(0.009)
Marital status (reference: single)						
Married	0.094***	0.106***	0.095***	0.063***	0.060***	0.052***
	(0.008)	(0.012)	(0.010)	(0.008)	(0.007)	(0.009)
Widowed	0.031	0.010	0.003	0.015	0.011	-0.005
	(0.020)	(0.028)	(0.022)	(0.016)	(0.022)	(0.023)
Divorced	0.024	-0.026	-0.014	-0.004	0.009	0.020
	(0.022)	(0.061)	(0.024)	(0.031)	(0.011)	(0.029)
Human capital						
Age	0.022***	0.025***	0.022***	0.020***	0.014***	0.014***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)	(0.002)
Age squared	-0.000***	-0.000***	-0.000***	-0.000***	-0.000***	-0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Education (reference: no education)						
Primary and secondary	0.155***	0.163***	0.144***	0.138***	0.147***	0.166***
	(0.039)	(0.037)	(0.043)	(0.047)	(0.037)	(0.053)
High school	0.250***	0.258***	0.231***	0.229***	0.231***	0.230***

		(0.039)	(0.038)	(0.043)	(0.047)	(0.037)	(0.052)
	University and above	0.533***	0.515***	0.501***	0.487***	0.433***	0.470***
		(0.040)	(0.039)	(0.044)	(0.047)	(0.037)	(0.053)
Occupation		Yes	Yes	Yes	Yes	Yes	Yes
Job characteristics							
Industry (reference: agriculture)							
	Manufacturing	0.014	-0.144***	-0.064***	0.068***	0.140***	0.162***
		(0.011)	(0.018)	(0.015)	(0.011)	(0.011)	(0.016)
	Trade	0.108***	0.147***	0.123***	0.148***	0.136***	0.064***
		(0.010)	(0.014)	(0.012)	(0.009)	(0.011)	(0.013)
	Services	-0.186***	-0.397***	-0.329***	-0.145***	-0.004	0.065***
		(0.010)	(0.015)	(0.012)	(0.010)	(0.011)	(0.014)
Public sector		0.253***	0.240***	0.302***	0.292***	0.264***	0.202***
		(0.010)	(0.018)	(0.013)	(0.010)	(0.008)	(0.011)
Geographical characteristics							
Rural area		-0.075***	-0.070***	-0.070***	-0.062***	-0.067***	-0.074***
		(0.007)	(0.010)	(0.008)	(0.006)	(0.006)	(0.008)
Region		Yes	Yes	Yes	Yes	Yes	Yes
Constant		3.124***	2.562***	3.050***	3.190***	3.517***	3.708***
		(0.065)	(0.111)	(0.085)	(0.062)	(0.060)	(0.099)
Observations		33,392	33,392	33,392	33,392	33,392	33,392
R-squared		0.566					

*Source: The Philippines' LFS 2009, Vietnam's LFS 2007. Authors' calculation. Robust standard errors in parentheses; ***p<0.01, **p<0.05, *p<0.1.*

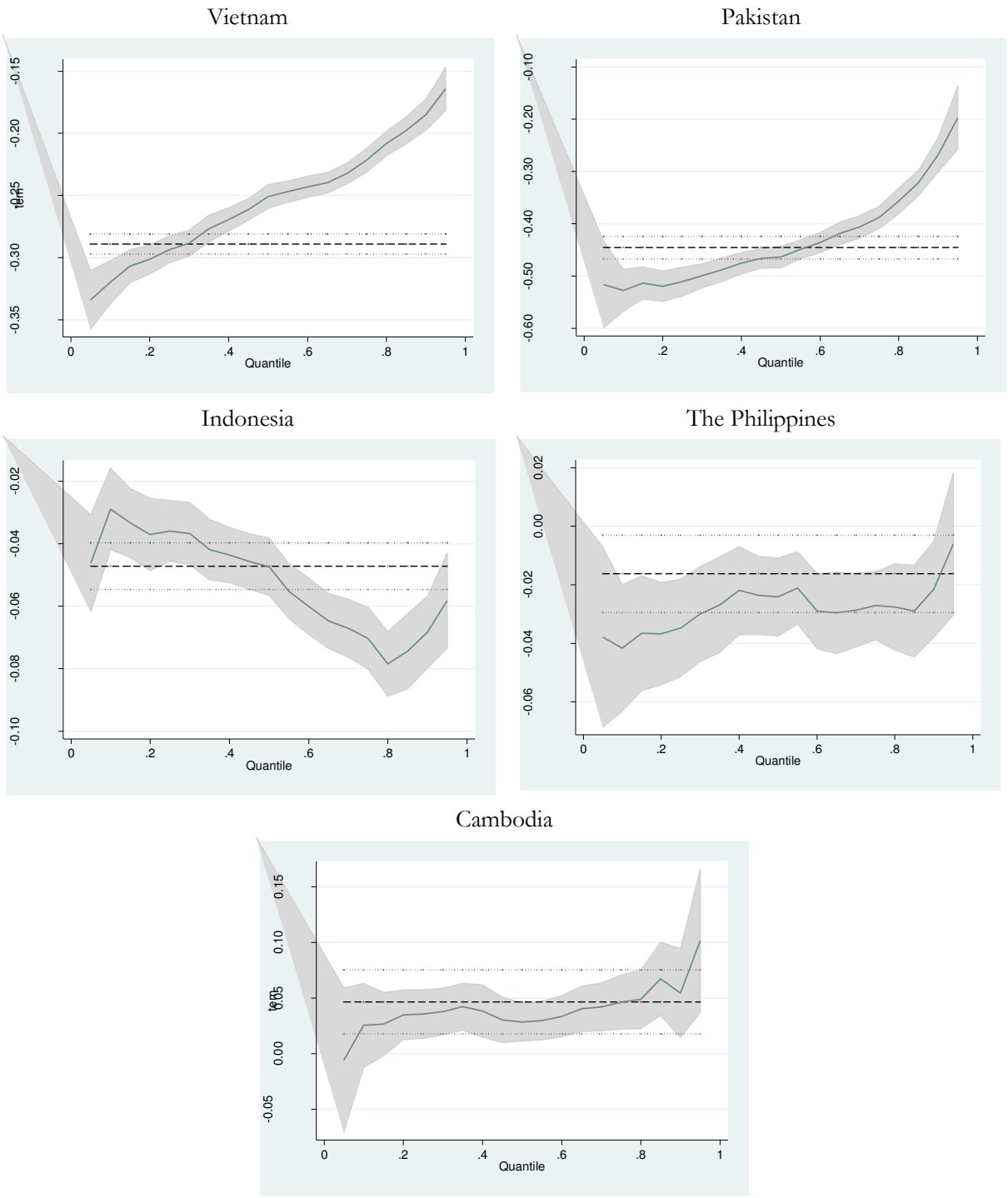
Table 10. Temporary-Permanent Workers' Wage Gap: Extended OLS and Quantile Regression – Vietnam

VARIABLES	(1) OLS	(2) QR .10	(3) QR .25	(4) QR .50	(5) QR .75	(6) QR .90
Temporary Employee	-0.289***	-0.320***	-0.293***	-0.251***	-0.221***	-0.185***
	(0.005)	(0.007)	(0.005)	(0.004)	(0.005)	(0.007)
Individual characteristics						
Gender Female	-0.156***	-0.150***	-0.146***	-0.157***	-0.164***	-0.171***
	(0.003)	(0.005)	(0.004)	(0.003)	(0.004)	(0.005)
Ethnic minority	0.001	-0.033***	-0.003	0.003	0.013*	0.025***
	(0.006)	(0.011)	(0.008)	(0.006)	(0.007)	(0.009)
Head of household	0.022***	0.035***	0.032***	0.025***	0.014***	0.016***
	(0.004)	(0.006)	(0.005)	(0.004)	(0.004)	(0.005)
Marital status (reference: single)						
Married	0.020***	0.038***	0.031***	0.019***	0.011**	0.005
	(0.004)	(0.007)	(0.005)	(0.004)	(0.005)	(0.007)
Widowed	-0.055***	-0.010	-0.068***	-0.063***	-0.050***	-0.031
	(0.013)	(0.011)	(0.017)	(0.014)	(0.017)	(0.022)
Divorced	-0.007	0.032	-0.011	-0.017	-0.005	-0.019
	(0.011)	(0.021)	(0.014)	(0.013)	(0.012)	(0.017)
Human capital						
Age	0.043***	0.050***	0.044***	0.037***	0.035***	0.033***
	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.002)
Age squared	-0.000***	-0.001***	-0.001***	-0.000***	-0.000***	-0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Education (reference: no education)						
Primary and secondary	0.049***	0.016	0.061***	0.062***	0.057***	0.070***
	(0.006)	(0.012)	(0.007)	(0.007)	(0.008)	(0.011)
High school	0.204***	0.162***	0.208***	0.215***	0.209***	0.214***
	(0.007)	(0.013)	(0.008)	(0.007)	(0.009)	(0.012)
University and above	0.550***	0.615***	0.592***	0.531***	0.470***	0.477***
	(0.008)	(0.013)	(0.009)	(0.008)	(0.009)	(0.013)

Job characteristics						
Industry (reference: agriculture)						
Manufacturing	0.028***	0.112***	0.053***	0.015**	-0.011	-0.072***
	(0.007)	(0.011)	(0.009)	(0.006)	(0.008)	(0.012)
Trade	0.049***	0.102***	0.057***	0.028***	0.018*	-0.019
	(0.008)	(0.013)	(0.010)	(0.008)	(0.009)	(0.014)
Services	0.045***	0.048***	0.030***	0.042***	0.061***	0.035***
	(0.007)	(0.012)	(0.009)	(0.007)	(0.009)	(0.013)
Institutional sector (reference: domestic private sector)						
Public sector	0.001	-0.031***	-0.018**	0.006	0.015	-0.007
	(0.007)	(0.010)	(0.007)	(0.007)	(0.011)	(0.018)
Foreign sector	-0.044***	-0.112***	-0.053***	0.013*	0.035***	0.009
	(0.008)	(0.010)	(0.008)	(0.007)	(0.011)	(0.018)
Informal sector	0.003	0.031***	0.033***	0.011**	-0.018***	-0.038***
	(0.005)	(0.009)	(0.006)	(0.005)	(0.006)	(0.007)
Business size	Yes	Yes	Yes	Yes	Yes	Yes
Geographical characteristics						
Rural area	-0.056***	-0.062***	-0.038***	-0.037***	-0.044***	-0.060***
	(0.004)	(0.006)	(0.004)	(0.004)	(0.004)	(0.005)
Region	Yes	Yes	Yes	Yes	Yes	Yes
Constant	0.876***	0.269***	0.577***	0.914***	1.225***	1.564***
	(0.022)	(0.036)	(0.027)	(0.021)	(0.025)	(0.036)
Observations	117,915	117,915	117,915	117,915	117,915	117,915
R-squared	0.372					

*Source: Vietnam's LFS 2007. Authors' calculation. Robust standard errors in parentheses; ***p<0.01, **p<0.05, *p<0.1.*

Figure 1. Temporary-Permanent Workers' Wage Differentials in the Selected Asian Countries, at Mean and Different Quantiles. Extended Set of Explanatory Variables.



Source: Cambodia's Labor Force and Child Labor Survey 2012, Indonesia's LFS 2007-2008, Pakistan's LFS 2008-2009, the Philippines' LFS 2009, Vietnam's LFS 2007. Author's calculation.

4.3. How we define 'temporary employment' matters

The cross-country analyses of temporary employment's wage gap raise the issue of comparability. As mentioned in section 3.2, 'temporary employment' is not homogeneously defined in the five national surveys (**Table 1**). With Vietnamese and Pakistani data, permanent jobs are characterized by both unlimited term and a written contract, leaving permanent jobs without a written contract to also be considered as temporary employment. Meanwhile, Cambodia's survey provides detailed information on both duration and type of contract, allowing to include 'unlimited duration oral contracts' in the permanent employment category. Thus, temporary employment in Cambodia includes only 'limited/unspecified duration written contracts,' and 'limited/unspecified duration oral contracts.'

This section aims to assess how the estimates on temporary workers' wage differentials would change in response to the modification in temporary employment's definition. Certainly, the previously stated main results are based on the most accurate and appropriate ways to define temporary employment given available survey information. First, to make the concept of temporary employment in Cambodia closer to that of Vietnam and Pakistan, a job would be considered permanent only if it has unlimited duration and a written contract. Second, temporary workers in Vietnam are re-defined as those with a fixed term or undefined contract but precarious duration *and* without social security coverage. This new definition narrows the scope of temporary employment in Vietnam, to be more detailed, makes it an even more 'inferior' job category. Third, in Pakistan, only jobs with contracts equal or shorter than one year or without contract/agreement would be considered as temporary employment. This one year threshold comes from Pakistani regulation that imposes the maximum duration of FTCs, including renewals, to be one year and less (ILO, 2015d). In addition, an alternative question is used to determine temporary employment (question 5.8 instead of question 7.1). According to this question, permanent wage workers are defined as 'regular paid employee with fixed wage,' and temporary workers as 'casual paid employee,' 'paid workers by piece rate or work performed,' or 'paid non-family apprentice.' The estimation results are presented in **Table 11**.

On average, the wage premium in Cambodia fades away once the new definition of temporary employment is adopted (OLS coefficient is small and no longer significant), but it is still observed for certain percentiles (e.g., 50^{th}, 75^{th}, 90^{th} quantiles). The mean estimate turns insignificant possibly because this modified definition has enlarged the scope of temporary jobs by imposing more 'superior' criteria on permanent employment, making the former less advantageous relative to the latter. Interestingly, the estimated wage differential for the 10^{th} quantile in the benchmark model has turned into negative (minus 5.7 percent), implying that the impact of 'written contract' criterion is the strongest on workers at the bottom of the pay ladder.

By contrast, the results in Vietnam do not change much when social security is taken into account. This could be explained by the fact that up to three quarters of the initially defined temporary workers already lacked social security. Adding this criterion to temporary employment thus does not cause any significant change in the estimated wage gaps.

Likewise, imposing one year duration threshold on temporary jobs slightly reduces the estimated wage penalties in Pakistan (from 45 percent to 38 percent in the benchmark model and

from 26 percent to 18 percent in the extended model). In fact, limited term contracts longer than one year constituted merely under six percent of all wage jobs, so excluding them from temporary employment does not change the estimated results very much. By contrast, when temporary employment is re-defined using question 5.8, the estimated wage gaps drop enormously (from 45 percent to five percent in the benchmark model and from 26 percent to 10 percent in the extended model). Actually, the concept of permanent wage workers defined using this question coincides with regular wage workers. According to Ahmad and Ahmad (2006), in Pakistan, "a 'regular' worker is enrolled on the labor register of the firm, has job security, is paid wages even when there is no work and receives benefits and facilities such as social security, medical care, education cess, pension, housing, transport, etc." The Resolution concerning the International Classification of Status in Employment (ICSE-93) also states that: "It is also recommended to distinguish *regular employees* from other 'employees with stable contracts' on the basis of the extent to which these contracts oblige the 'employer' to pay regular social security contributions and/or are subject to national labor legislation."[86] This definition is closer to formal wage employment than to permanent wage employment, without any explicit reference to duration. Meanwhile, the original definition of permanent employment (i.e., the one used for the main analyses) implies both permanency and formality ('permanent/pensionable job'). Data show that only 45 percent of permanent workers defined with question 5.8 could be considered as permanent workers defined as 'permanent/pensionable job.' Therefore, it is reasonable that the temporary-permanent wage penalty is more narrow than before.

To sum up, the robustness checks have illustrated that definitions of temporary employment do matter to some extent, particularly when cross-country comparisons are to be made. In cases of Indonesia and the Philippines, there is no second way to define temporary employment with given questionnaires. Pakistani data, by contrast, offer several alternative definitions, which might lead to more or less different results. Although the chosen definitions in this paper (section 3.2) offer the as much compatibility as possible with the international guidance given each country's dataset, it is recommended to keep in mind the potential bias due to discrepancies in the definitions of temporary employment in the selected countries. These results should be interpreted given the exact scope of the so-called 'temporary employment' in different country contexts.

[86] Source: http://www.ilo.org/global/statistics-and-databases/statistics-overview-and-topics/status-in-employment/current-guidelines/lang--en/index.htm

Table 11. Temporary-Permanent Workers' Wage Gap: OLS and Quantile Regression – Modified Definitions of Temporary Employment

	Benchmark model						Extended model					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
VARIABLES	OLS	QR .10	QR .25	QR .50	QR .75	QR .90	OLS	QR .10	QR .25	QR .50	QR .75	QR .90
Cambodia	0.025	-0.057***	-0.011	0.025**	0.051***	0.061**	0.037	0.022	0.004	0.020	0.045***	0.042**
	(0.023)	(0.022)	(0.014)	(0.012)	(0.017)	(0.028)	(0.023)	(0.020)	(0.012)	(0.013)	(0.015)	(0.021)
Pakistan (1)	-0.381***	-0.401***	-0.395***	-0.388***	-0.358***	-0.291***	-0.178***	-0.189***	-0.188***	-0.180***	-0.144***	-0.114***
	(0.011)	(0.018)	(0.011)	(0.009)	(0.010)	(0.013)	(0.015)	(0.020)	(0.014)	(0.010)	(0.011)	(0.012)
Pakistan (2)	-0.052***	-0.062***	-0.067***	-0.059***	-0.030***	0.024*	0.103***	0.055***	0.101***	0.121***	0.146***	0.143***
	(0.010)	(0.017)	(0.011)	(0.009)	(0.010)	(0.013)	(0.011)	(0.016)	(0.011)	(0.009)	(0.010)	(0.014)
Vietnam	-0.302***	-0.303***	-0.277***	-0.275***	-0.279***	-0.291***	-0.290***	-0.301***	-0.277***	-0.252***	-0.241***	-0.237***
	(0.004)	(0.007)	(0.005)	(0.004)	(0.004)	(0.006)	(0.005)	(0.008)	(0.006)	(0.005)	(0.006)	(0.007)

*Source: Cambodia's Labor Force and Child Labor Survey 2012, Vietnam's LFS 2007, the Philippines' LFS 2009, Pakistan's LFS 2008-2009, Indonesia's LFS 2007-2008. Authors' calculation. Robust standard errors in parentheses; ***p<0.01, **p<0.05, *p<0.1.*

Note: Pakistan (1) defines temporary workers as those with contracts equal or shorter than one year (question 7.1); Pakistan (2) defines temporary workers as regular paid employees with fixed wage (question 5.8).

4.4. Discussion and interpretation of the estimation results

Diverse estimated results appear despite the fact that all studied countries share a number of labor market characteristics such as minimum wage law and labor unions. Another study by Mertens et al. (2007) also points out different experiences of wage penalty associated with fixed-term jobs between Germany and Spain, although they are both considered as 'rigid' economies. In fact, there are many other institutional as well as socioeconomic factors being at work. This section attempts to provide an interpretation and discussion of the results by taking into account these relevant factors as much as possible.

Table 12 and **Figure 3** (Appendix) provide some macroeconomic indicators across selected countries. First, there seems to be a relationship between the distribution pattern of temporary employment's wage gap in a country and their current stage of development, proxied by the GDP per capita (except the very special case of Cambodia). According to the literature explaining the *sticky floor* or *glass ceiling* wage gaps in gender context, *glass ceiling* effect is almost present in high income-low inequality regions/countries (Europe, North America, and Australia) while in developing countries, it seems that *sticky floor* effect is the norm (Albrecht et al., 2003; Arulampalam et al., 2007; Fang & Sakellariou, 2015). Our results on temporary-permanent workers' wage gaps are consistent with this phenomenon. Indeed, *sticky floor* effect appears in the least developed among the studied countries (Vietnam, Pakistan) where the wage penalty is also the most severe. By contrast, *glass ceiling* phenomenon exhibits in Indonesia, the most developed among the studied countries. In the Philippines – a medium income country, the wage gap is stable across the wage distribution.

Second, the wage differentials between temporary and permanent workers tend to be related with the unemployment rate (*'unemp.'*) in the economy. For example, in Cambodia, an extremely low unemployment rate coupled with a high demand for low skilled labors and FTC workers in garment industry might induce Cambodian temporary workers to be offered with competitive wage rates compared to their permanent counterparts. Meanwhile, in the Philippines and Indonesia, the unemployment rate is relatively high, leading to limited wage premiums for permanent workers. Other macroeconomic indicators such as growth rate, inflation rate, labor force size or human development index (HDI) do not exhibit an evident relationship with the temporary-permanent workers' wage differentials in the countries.

For further interpretation of the wage gaps, more country-specific contexts should be investigated. Besides, separate estimations of the wage differentials for men and women are performed and reported in **Table 13** and **Figure 2**.

Table 12. Macroeconomic Conditions in Selected Asian Countries: A Comparative View

Country	Year	GDP per capita, PPP	Growth	Inflation	Unemp.	Labor force	HDI
Cambodia	2007	2323.89	10.21	7.7	0.6	7.26	
	2008	2442.92	6.69	25	0.3	7.5	0.564
	2009	2408.62	0.09	-0.7	0.1	7.7	
	2010	2513.29	5.96	4	0.4	7.9	0.571
	2011	2648.65	7.07	5.5	0.3	8.08	0.575
	2012	2795.17	7.26	2.9	0.2	8.26	0.579
	2013	2955.17	7.48	2.9	0.3	8.45	0.584
	2014	3112.63	7.07	3.9	0.4	8.62	0.555
Indonesia	2007	7472.81	6.35	6.4	9.1	111.24	
	2008	7819.07	6.01	9.8	8.4	113.03	0.654
	2009	8074.5	4.63	4.8	7.9	115.05	
	2010	8465.3	6.22	5.1	7.1	116.5	0.671
	2011	8870.28	6.17	5.4	6.6	118.52	0.678
	2012	9282.71	6.03	4.3	6.1	120.43	0.681
	2013	9672.6	5.56	6.4	6.3	122.13	0.684
	2014	10031.32	5.02	6.4	6.2	124.06	0.684
Pakistan	2007	4303.5	4.83	7.6	5.1	53.36	
	2008	4287.38	1.7	20.3	5	54.67	0.536
	2009	4318.13	2.83	13.6	5.2	56.65	
	2010	4296.61	1.61	13.9	5.1	58.6	0.526
	2011	4322.53	2.75	11.9	5.1	60.16	0.531
	2012	4380.24	3.51	9.7	5.1	61.82	0.535
	2013	4475.65	4.37	7.7	5.1	63.65	0.537
	2014	4590.15	4.74	7.2	5.2	65.36	0.538
Philippines	2007	5200.1	6.62	2.9	7.4	36.09	
	2008	5336.18	4.15	8.3	7.3	37.34	0.648
	2009	5318.26	1.15	4.2	7.5	38.54	
	2010	5638.21	7.63	3.8	7.3	39.96	0.651
	2011	5754.11	3.66	4.6	7	41.14	0.652
	2012	6041.78	6.68	3.2	7	42.03	0.656
	2013	6365	7.06	3	7.1	42.92	0.66
	2014	6648.55	6.13	4.1	7.1	43.81	0.668
Vietnam	2007	3907.27	7.13	8.3	2.3	48.03	
	2008	4084.82	5.66	23.1	2.4	49.01	0.617
	2009	4259.96	5.4	7.1	2.6	49.97	
	2010	4486.26	6.42	8.9	2.6	50.89	0.629
	2011	4715.89	6.24	18.7	2	51.77	0.632
	2012	4910.31	5.25	9.1	1.8	52.66	0.635
	2013	5121.74	5.42	6.6	2.2	53.47	0.638
	2014	5370.21	5.98	4.1	2.3	54.21	0.666

Source: theglobaleconomy.com, the World Bank.

Indonesia

The *glass ceiling* effect observed with Indonesian data, for both men and women, might be related to the presence of equal opportunity and anti-discrimination laws. The Manpower Protection Act in 2003,[87] for example, emphasizes the protection of workers, defined as "every person who works for a wage or other forms of remuneration," which is intended to "secure the implementation of equal opportunity and equal treatment without discrimination on whatever basis in order to realize the welfare of workers." Indonesia also introduced a three-pillar social security system, including a publicly-funded social assistance, which benefit the old and poor. The National Social Security Law (2004)[88] mandates the extension of social security coverage to the whole population, applying non-contributory schemes for the poor. Indonesia also strives to provide forms of protection and assistance to the disadvantaged groups through the programs such as National People Empowerment Program and Conditional Cash Transfer scheme.

Vietnam and Pakistan

The *sticky floor* effect found in Vietnam and Pakistan could be possibly linked to the predominance of the informal economy and the difficulty to include temporary workers in trade unions. First, as far as the informal economy's concerned, in 2007, nearly half of non-farm jobs in Vietnam located in the informal sector (Cling et al., 2010), and 75 percent of temporary jobs were also informal. This sector is generally made up of low-income jobs with precarious working conditions, lacking protection such as labor contract, minimum wage, social security and so on (Cling et al., 2014). Effectively, Nguyen et al. (2013) report that informal wage workers earn, on average, 23 percent lower than their formal counterparts. The gap shrinks to 11 percent if unobserved individual characteristics are taken into account.

Similarly, Pakistan also witnesses a prevalence of the informal sector. In 2009, this sector accounted for about 70 percent of employment in main job (ILO, 2015b). Although informal sector employees suffer longer working hours than their formal sector counterparts (60 hours versus 45 hours), the former are poorly remunerated and receive smaller fringe benefits (Kemal & Mahmood, 1998). Overall, the informal economy is characterized by poor working conditions, particularly for female employees who are not covered under legal protection. As reported in ILO (2015b), one of the main challenges of the country is the "weak application of national labor legislation and regulations to increasingly formalize the large and growing informal economy." In summary, although Vietnam and Pakistan do have a minimum wage law, it does not benefit the numerous informal temporary wage workers. As a result, those disadvantaged workers are left unprotected, vulnerable and highly discriminated against in terms of payment, particularly in the lower tiers of the wage distribution.

Second, Vietnamese and Pakistani trade unions find it difficult to protect temporary workers. As previously mentioned, the VGCL is the only trade union in Vietnam, and it is difficult for trade unions to include and support temporary workers (Tung, 2013; as cited in Serrano et al., 2014). Moreover, whether VGCL plays a similar role as a typical trade union in developed countries

[87] Source: ilo.org/dyn/travail/docs/760/Indonesian%20Labor%20Law%20-%20Act%2013%20of%202003.pdf

[88] Source: ituc-csi.org/the-reform-of-social-security-in?lang=en

remains uncertain.[89] Likewise, ILO (2015b) indicates that in Pakistan, the organized representation of employees and employers mainly exists in the formal economy but remains at a low level.

Table 13 shows that while Vietnamese men and women face nearly similar levels of temporary workers' wage penalty, in Pakistan, temporary female workers are punished twice as much as their male counterparts. It could be the case that either temporary female workers earn too little or permanent female workers earn very much, but the former possibility seems to be more reasonable. Besides, both sex groups experience a *sticky floor* effect in the two countries. These suggest that the temporary-permanent work remuneration penalty is the most acute among Pakistani women who hold a low-wage temporary job.

The Philippines

In the Philippines, the temporary job's wage penalty is not only modest but also stable across the wage distribution. Such a high remuneration equality might be related to their labor market institution and regulation as well as trade union's strategy. The Filipino Labor Code defines legitimate contracting and subcontracting with a number of requirements, including the rights to the contractor's employees such as safe and healthy working conditions, labor standards (e.g., service incentive leave, rest days, overtime pay, holiday pay, 13th month pay and separation pay), social security and welfare benefits, the right to self-organization, collective bargaining and security of tenure. In addition, the Philippines adopted the Social Protection Operational Framework and Strategy (SPOFS) in 2012, which particularly aims at increasing employment opportunities and enhancing protection of workers' rights and welfare.[90] Moreover, Filipino trade unions have also employed complementary strategies at the enterprise, industry and national levels to bring more protection to non-regular workers (Serrano et al., 2014).

Interestingly, when men and women are investigated separately, opposite phenomena are observed. On average, men face a larger temporary wage penalty than do women (4.7 percent versus 2.4 percent). While no distributional effect of the penalty along the wage distribution can be assured for the former due to large confidence intervals, the latter group exhibit a clear *sticky floor* effect. Temporary female workers earn less than their permanent counterpart by 7.1 percent at the 10th quantile, but the gap gradually diminishes and even turns into a premium in the upper tier of the wage distribution (5.7 percent at the 90th quantile). These results may imply that anti-discrimination policies are more effective in protecting males and high wage female earners.

[89] The VGCL itself has declared that "Under the leadership of the Communist Party of Vietnam, ever since its founding, the Vietnamese trade unions have always been loyal to the interests of the working class and nation; and organized and mobilized workers, office employees and other working people to pioneer in the struggle for the independence and freedom of the homeland and the lawful and legitimate interests of the working people." (Source: factsanddetails.com/southeast-asia/Vietnam/sub5_9g/entry-3474.html)

According to a survey by ITUC, in practice, Vietnam is characterized by "no freedom of association," "internet under surveillance to prevent calls for strike action," "collective bargaining restricted," "decree makes suppression of strikes even easier," "labor right activists detained and beaten," etc. (Source: http://survey.ituc-csi.org/Vietnam.html?lang=en#tabs-3)

[90] Although SPOFS was adopted after the Philippines' LFS 2009 survey was conducted, this possibly reflects their early efforts to reinforce social protection.

Cambodia

Cambodia is the only country among the list with a wage advantage enjoyed by temporary employees. Recall that, based on Cambodia Labor Force and Child Labor Survey 2012, this study defines temporary employment as all wage jobs that have either 'limited duration' or 'unspecified duration' or 'don't know.' Meanwhile, the Cambodian Labor Law indicates that "Fixed-term contracts may have an unspecified finishing date when they are drawn up for replacing a worker who is temporarily absent; seasonal work; occasional periods of extra work or a non-customary activity of the enterprise." Evidently, what is referred as temporary employment in this paper mostly coincides with these FTC jobs. The temporary work's wage premium is, therefore, very likely to be relevant to the particularity of Cambodia's FTC jobs.

On the one hand, FTC jobs in Cambodia are characterized by a number of 'superior' characteristics. In 2012, over half of FTC workers in Cambodia were employed under a written contract while the rate for the whole wage workers was only 29 percent. The proportion of formal jobs among FTC workers (20 percent) was greater than the proportion of formal jobs among all wage workers (14 percent). Moreover, a majority of FTC lasted 12 months or more (73 percent), making them very close to a long-term stable employment. In Cambodia, FTC jobs are also authorized for permanent tasks.[91] Cambodia is the only country in which both public and private sector firms employed more temporary than regular wage workers, and the use of temporary employees was even more prevalent in the public sector.

On the other hand, Cambodian workers under a limited duration contract enjoy fewer rights and benefits than their unlimited-term counterparts – including paid annual leave, seniority rights and maternity leave (ITUC, 2014). It is also much easier to fire FTC workers than permanent ones (ibid). In order to compensate for these disadvantages, FTCs might offer a more competitive wage rate (Compensating Differentials theory). Anecdotal evidence also suggests that workers in Cambodia receive a five percent payment at the end of the contract, rendering such contracts attractive to workers, and hence possibly leading to positive self-selection of workers in these types of contracts. Indeed, the best educated workers, including those having completed a university education, as well as other high skilled workers such as managers and professionals engaged the most in temporary work. This compensatory mechanism is likely to explain the wage premium attributed to limited-term written contract workers in this country.[92]

The temporary workers' wage premium in Cambodia could also be linked to the high demand for low skilled labors in Cambodian labor market. Cambodia's growth has been mostly based on low productivity sectors that do not require high levels of education.[93] For instance, the main driving force of Cambodia's economy – the garment industry – focuses on low-skilled, labor-intensive activities; and about 90 percent of the labor force in this sector are female workers with

[91] Source: ilo.org/wcmsp5/groups/public/---ed_protect/---protrav/---travail/documents/publication/wcms_357403.pdf

[92] This hypothesis should be yet taken with precaution. Warren and Robertson (2011) find a robust positive relationship between wages and working conditions in Cambodian garment factories, suggesting that efficiency wage theory more accurately explains the behavior of these firms than compensating wage theory.

[93] For more information, visit: http://www.khmertimeskh.com/news/27088/education-must-be-rewarded/

minimum skills (USAID, 2006). ITUC (2014) asserts that many garment factories have built an entire workforce of workers hired on repeatedly-renewed short-term FTCs and this trend is increasingly popular.[94] Thus, the young educated Cambodians, who prefer to pursue careers in areas such as finance, business and law, have little interest in the garment industry with limited career potential. According to USAID (2006), in areas close to Phnom Penh with a high density of factories and a large demand for workers, the recruitment becomes so difficult that investors are obliged to offer a higher salary and additional benefits apart from the minimum wage.[95] Differently speaking, an important share of FTC work – those in the textile, footwear and garment industry – could possibly receive higher wages than an average permanent wage worker does. Separate quantile regression results by gender reveal that such a wage premium actually takes place in the middle range of the wage ladder for men and in the upper tail of the wage distribution for women.

[94] Our current data analysis shows that 65 percent of wage workers in garment and related pattern-makers and cutters were of FTC.

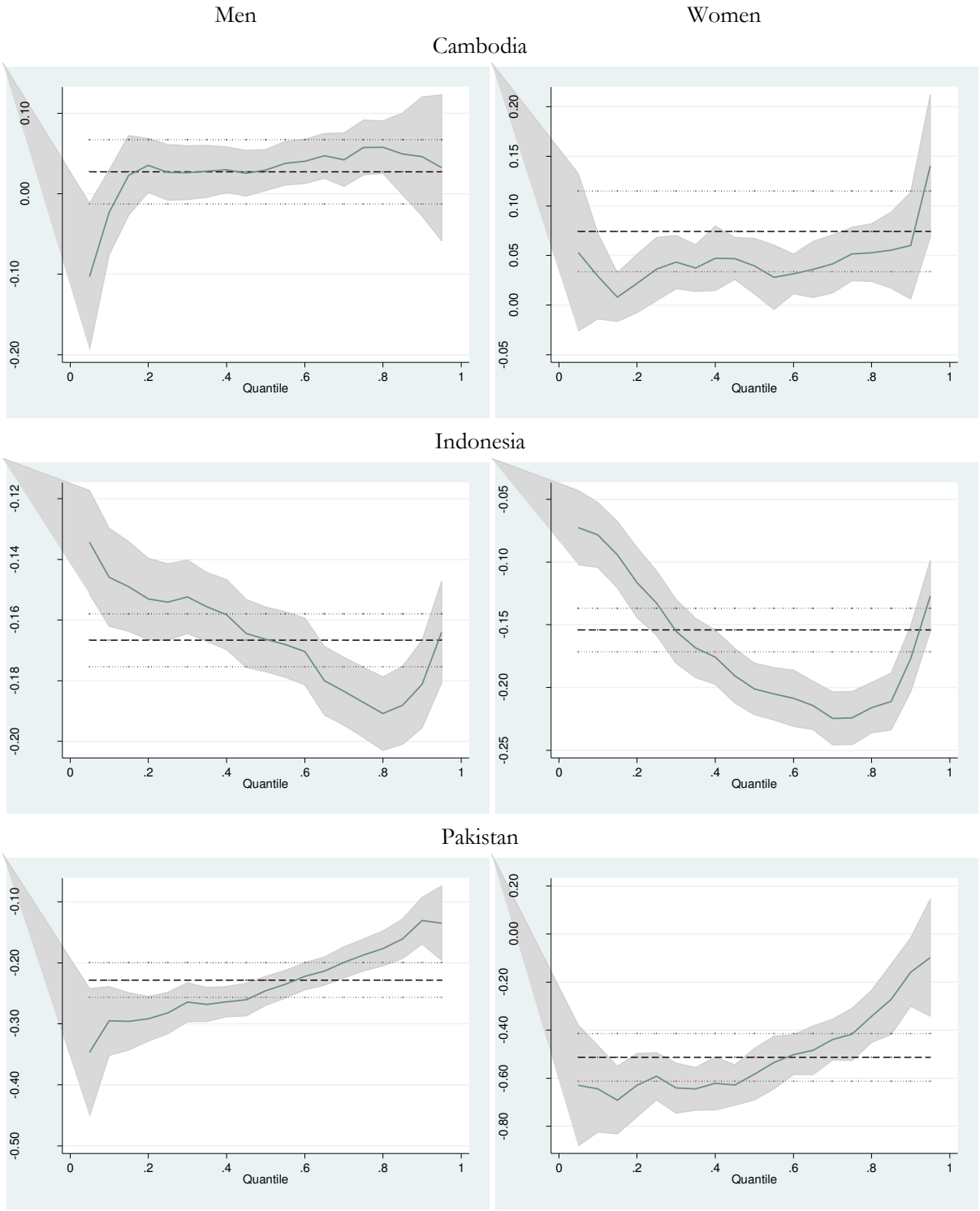
[95] The textile, footwear and garment industry is the only sector covered with a minimum wage rate. Better Factories Cambodia reports show a very high compliance rate of minimum wage payment in this sector, at 99 percent for regular workers and 89 per cent for casual workers (ILO, 2012).

Table 13. Temporary-Permanent Workers' Wage Gap: OLS and Quantile Regression, Extended Models by Gender

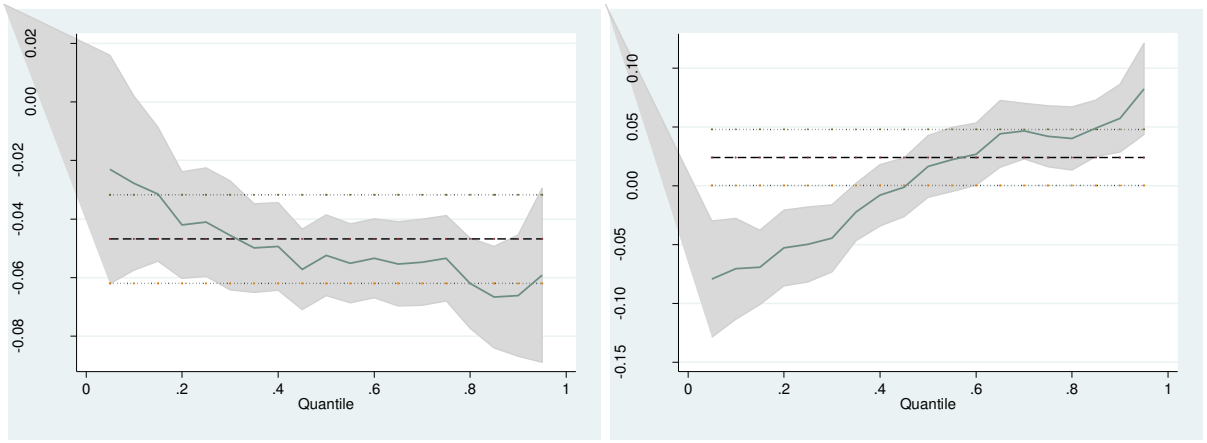
	Men						Women					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	OLS	QR .10	QR .25	QR .50	QR .75	QR .90	OLS	QR .10	QR .25	QR .50	QR .75	QR .90
Cambodia	0.027	-0.023	0.027**	0.029***	0.058***	0.046	0.074***	0.029	0.036	0.039	0.052	0.060***
	(0.020)	(0.000)	(0.012)	(0.011)	(0.012)	(0.000)	(0.021)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Indonesia	-0.167***	-0.146***	-0.154***	-0.166***	-0.187***	-0.181***	-0.154***	-0.078***	-0.133***	-0.201***	-0.224***	-0.177***
	(0.004)	(0.008)	(0.006)	(0.005)	(0.006)	(0.007)	(0.009)	(0.011)	(0.011)	(0.010)	(0.011)	(0.014)
Pakistan	-0.228***	-0.295***	-0.282***	-0.246***	-0.187***	-0.131***	-0.514***	-0.644	-0.592	-0.584***	-0.417***	-0.159***
	(0.017)	(0.021)	(0.015)	(0.011)	(0.013)	(0.016)	(0.050)	(0.000)	(0.000)	(0.043)	(0.046)	(0.046)
Philippines	-0.047***	-0.028**	-0.041***	-0.052***	-0.053***	-0.066***	0.024*	-0.071***	-0.050***	0.016	0.042***	0.057***
	(0.008)	(0.012)	(0.008)	(0.006)	(0.006)	(0.009)	(0.013)	(0.016)	(0.012)	(0.000)	(0.011)	(0.014)
Vietnam	-0.285***	-0.326***	-0.298***	-0.246***	-0.214***	-0.170***	-0.272***	-0.311***	-0.288***	-0.239***	-0.210***	-0.171***
	(0.007)	(0.010)	(0.007)	(0.006)	(0.007)	(0.009)	(0.007)	(0.009)	(0.007)	(0.006)	(0.007)	(0.009)

Source: Cambodia's Labor Force and Child Labor Survey 2012, Indonesia's LFS 2007-2008, Pakistan's LFS 2008-2009, the Philippines' LFS 2009, Vietnam's LFS 2007. Authors' calculation.
*Robust standard errors in parentheses; ***p<0.01, **p<0.05, *p<0.1.*

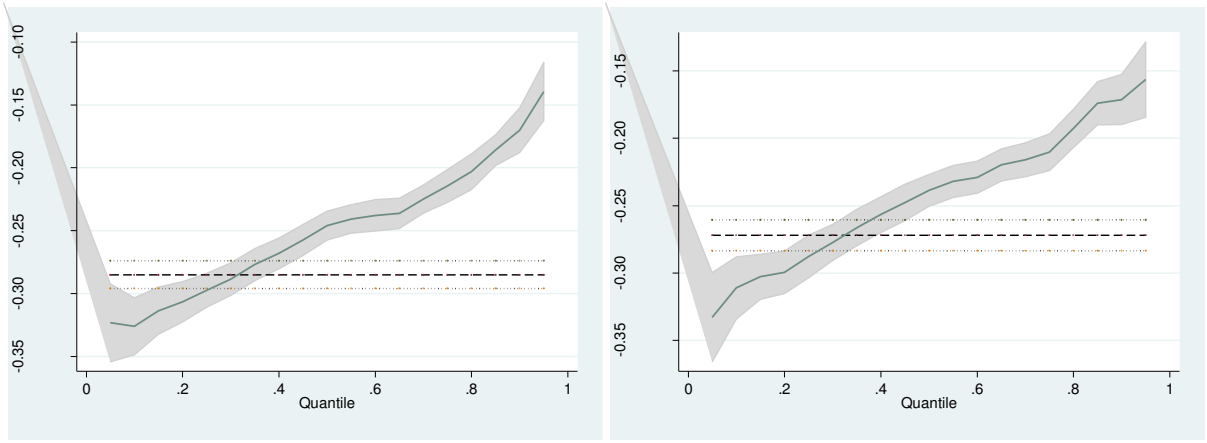
Figure 2. Temporary-Permanent Workers' Wage Differentials by Gender in the Selected Asian Countries, at Mean and Different Quantiles. Extended Set of Explanatory Variables.



The Philippines



Vietnam



Source: Cambodia's Labor Force and Child Labor Survey 2012, Indonesia's LFS 2007-2008, Pakistan's LFS 2008-2009, the Philippines' LFS 2009, Vietnam's LFS 2007. Author's calculation.

5. Concluding remarks

This study investigates the temporary-permanent workers' wage differentials in some selected Asian countries. Our review of literature has shown that though some Asian-country evidence on temporary employment exists, no systematic comparative research has been conducted focusing on this topic. There is, indeed, a scarcity of cross-nation statistical evidence on this employment form in Asia. Such a shortage of Asian cross-country statistics and studies on temporary employment could be attributed to the fact that this form of employment has been less of a concern as compared to informal employment, that there has been a lack of awareness of the phenomenon, or also because of lacking the data to study this phenomenon properly. This study manages to provide first-time-ever empirical evidence, shedding light on this neglected research area for some developing countries in Asia, including Cambodia, Indonesia, Pakistan, the Philippines and Vietnam.

In the theories, the wage gap associated with temporary employment is still inconclusive. On the one hand, Human Capital theory, Efficiency Wage theory, Labor Market Segmentation theory and Insider-Outsider model all indicate a wage penalty suffered by the temporary workers.

On the other hand, the theory of Equalizing Differences, or Compensatory Differences, predicts that temporary workers could receive wage premiums to compensate for undesirable non-wage characteristics associated with their jobs. Our analyses of temporary-permanent workers' wage gaps actually find supporting evidence for both wage penalty and wage premium. On average, temporary workers suffer from wage penalties of at least two percent in Indonesia, Pakistan, the Philippines and Vietnam, but benefit from a wage premium of five percent in Cambodia. Vietnam and Pakistan are the countries where temporary workers face the severest wage penalty. Temporary wage differentials highly depend on the relative position in the conditional wage distribution, except in the Philippines and Cambodia, where temporary workers encounter similar levels of wage penalty/premium at any positions in the pay ladder. Both *sticky floor* and *glass ceiling* effects emerge. The temporary employment's wage gap widens in the lower tiers of the wage distribution in Vietnam and Pakistan, suggesting that the penalty of being in temporary jobs could be more severe for disadvantaged workers, while the *glass ceiling* effect impedes Indonesian temporary workers from approaching high wages.

Different temporary-permanent workers' wage gaps in the five selected Asian countries are possibly related to their specific institutional and socioeconomic contexts. While smaller wage discrimination among low wage earners in Indonesia might be a result of their equal opportunities and anti-discrimination laws; the predominance of the informal economy and the difficulty to include temporary workers in trade unions in Vietnam and Pakistan are probably relevant to greater wage penalty suffered by workers at the bottom of the wage distribution. Meanwhile, the temporary employment's wage penalty is modest and stable along the wage distribution in the Philippines, where the labor market institution and regulation as well as trade unions highly protect non-regular workers. In the special case of Cambodia, temporary worker's remuneration advantage seems to be in line with Compensatory Wage theory. There is anecdotal evidence that Cambodian FTC workers receive a five percent payment at the end of the contract, in compensating for fewer rights and benefits such as paid annual leave. Particularly, due to limited career potential in garment industry, which is characterized by a broad use of repeatedly-renewed short-term FTCs, there is a large unmet demand for low skilled FTC workers, possibly leading to their higher wages than permanent workers'.

This general inequality between permanent and temporary workers is a challenge to be addressed, notably as temporary employment accounts for an even higher proportion in these developing Asian countries than in many other developed economies. Thus, reinforcement of the legal framework of labor contracts could be a central measure to improve employer-employee relationship and prevent the existence of loosely contractual engagement which leads to poor labor conditions and poorly paid jobs. However, as regulation is said to be working for the insiders but not the outsiders, social protection schemes should be delinked from employment status and extended to the informal sector usually on a non-contributory basis. Moreover, country-specific measures should be applied as different distribution patterns of wage penalties emerge. In Vietnam and Pakistan, as temporary workers possibly face rigorous pay discrimination when they first enter the labor market, assistance and protection policies should be focused on these individuals. Meanwhile, a further enhancement of labor protection such as minimum wage law might not be appropriate to narrow the pay gap at the upper tiers of the wage distribution in Indonesia.

Due to some limits of data sources, the current study leaves some issues to be further addressed. First, the heterogeneous definitions and classifications of temporary employment forms applied in LFS across selected countries hinder, to a certain extent, the comparability of the obtained results. For example, when the temporary employment concept in Cambodia is defined in another way, the temporary wage premium sustains in only several parts of the wage distribution. Fortunately, alternative definitions of temporary jobs in Vietnam and Pakistan (using the same question) do not result in significant changes in the estimated wage gaps. Second, further investigation focusing on demand side should be conducted to provide more evidence explaining the nexus between wage gaps and the wide use of temporary employment. It is necessary to understand if the importantly negative temporary wage gap (or permanent workers' wage premium) explains the popular use of temporary labor as a way to restrain wage bill, or if flexible use of employment through short-term contracts, casual jobs and so on is simply the norm in these developing countries.

References

Agrawal, T. (2013). Are There Glass-Ceiling and Sticky-Floor Effects in India? An Empirical Examination. *Oxford Development Studies*, *41*(3), 322-342.

Ahn, J. (2006). Nonstandard work in Japan and Korea – the origin of wage differentials. Non-standard Work-Draft.

Ahmad, I., & Ahmad, N. (2006). Pakistan National Studies in Employment Situations and Workers Protection. Retrieved from http://www.ilo.org/wcmsp5/groups/public/---ed_dialogue/---dialogue/documents/genericdocument/wcms_205373.pdf

Albrecht, J., Björklund, A., & Vroman, S. (2003). Is There a Glass Ceiling in Sweden?. *Journal Of Labor Economics*, *21*(1), 145-177.

Arulampalam, W., Booth, A., & Bryan, M. (2007). Is There a Glass Ceiling over Europe? Exploring the Gender Pay Gap across the Wage Distribution. *Industrial & Labor Relations Review*, *60*(2), 163-186.

Bentolila, S., Dolado, J., Franz, W., & Pissarides, C. (1994). Labour Flexibility and Wages: Lessons from Spain. *Economic Policy*, *9*(18), 55-99.

Bhandari, A., & Heshmati, A. (2008). Wage Inequality and Job Insecurity among Permanent and Contract Workers in India: Evidence from Organized Manufacturing Industries. *ICFAI Journal of Applied Economics, 7*(1), 80 – 111.

Blinder, A. (1973). Wage Discrimination: Reduced Form and Structural Estimates. *The Journal Of Human Resources*, *8*(4), 436-455.

Booth, A., Francesconi, M., & Frank, J. (2002). Temporary Jobs: Stepping Stones or Dead Ends?. *The Economic Journal, 122* (480), F189–F213.

Bosio, G. (2014). The Implications of Temporary Jobs on the Distribution of Wages in Italy : an Unconditional IVQTE Approach. *Labour: Review of Labour Economics and Industrial Relations, 28* (1, (3)), 64–86. IVQTE = unconditional instrumental variable quantile treatment effects.

Brown, C. (1980). Equalizing Differences in the Labor Market. *The Quarterly Journal Of Economics*, *94*(1), 113-134.

Christofides, L., Polycarpou, A., & Vrachimis, K. (2013). Gender Wage Gaps, 'Sticky Floors' and 'Glass Ceilings' in Europe. *Labour Economics*, *21*, 86-102.

Cling, J.-P., Nguyen, T. T. H., Nguyen, H. C., Phan, T. N. T., Razafindrakoto, M., & Roubaud, F. (2010). *The Informal Sector in Vietnam: A focus on Hanoi and Ho Chi Minh City*. Editions The Gioi.

Cling, J.-P., Razafindrakoto, M., & Roubaud, F. (2014). Segmentation and informality in Vietnam: A survey of the literature country case study on labour market segmentation. *Conditions of Work and Employment Series-ILO 52*.

Comi, S. & Brasseni, M. (2012). Are Temporary Workers Discriminated Against? Evidence from Europe. *The Manchester School, 80* (1), 28–50.

von Hippel, C., Mangum, S., Greenberger, D., Heneman, R., & Skoglind, J. (1997). Temporary Employment: Can Organizations and Employees Both Win?. *Academy Of Management Perspectives*, *11*(1), 93-104.

Dasgupta, S., Bhula-or, R., & Fakthong, T. (2015). Earnings differentials between formal and informal employment in Thailand. *ILO Asia-Pacific Working Paper Series*.

Debrah, Y., & Smith, I. (2002). *Globalization, Employment and the Workplace: Diverse Impacts*. Routledge.

Diamond, J. (2010). Non-standard Employment and Earnings in Japan. Retrieved from http://www.econ.hit-u.ac.jp/~diamond/Jess_Diamonds_Homepage/Research_files/ns%20wages.pdf

Doeringer, P., & Piore, M. (1971). *Internal Labor Markets and Manpower Analysis*. Lexington Books, 1st edition.

Doing Business (2009). *Doing Business 2010: Reforming through Difficult Time*. The World Bank, IFC and Palgrave MacMilan. Retrieved from http://www.doingbusiness.org/~/media/GIAWB/Doing%20Business/Documents/Annual-Reports/English/DB10-FullReport.pdf

Edgell, S. (2012). *The Sociology of Work: Continuity and Change in Paid and Unpaid Work, 2nd edition.* SAGE Publications Ltd.

Edin, P., & Richardson, K. (2002). Swimming with the Tide: Solidary Wage Policy and the Gender Earnings Gap. *Scandinavian Journal of Economics*, *104*(1), 49-67.

Esteban-Pretel, J., Nakajima, R., & Tanaka, R. (2011). Are Contingent Jobs Dead Ends or Stepping Stones to Regular Jobs? Evidence from a Structural Estimation. *Labour Economics*, *18*(4), 513-526.

Fang, Z., & Sakellariou, C. (2015). Glass Ceilings versus Sticky Floors: Evidence from Southeast Asia and an International Update. *Asian Economic Journal*, *29*(3), 215-242.

Fathollah, M., & Orhan, K. (2008). Selection bias in the estimated rate of return to education. http://www.freepatentsonline.com/article/Review-Business-Research/190699919.html

Gash, V., & McGinnity, F. (2007). Fixed-term Contracts—the New European Inequality? Comparing Men and Women in West Germany and France. *Socio-Economic Review*, *5*(3), 467-496.

Ge, Y., Li, H., & Zhang, J. (2011). Gender Earnings Gaps in Hong Kong: Empirical Evidence from across the Earnings Distribution in 2006. *China Economic Review*, *22*(1), 151-164.

Ghose, A. (2004). The Employment Challenge in India. *Economic and Political Weekly,* (48), 5106.

Güell, M. (2000). Fixed-Term Contracts and Unemployment: An Efficiency Wage Analysis. *Industrial Relations Section Working Paper*, (433).

Ha, B.-J., & Lee, S. (2013). Dual dimensions of non-regular work and SMEs in the Republic of Korea: Country case study on labor market segmentation. *ILO Employment Working Paper No. 148*.

Harrison, B., & Kelley, M. (1993). Outsourcing and the Search for `Flexibility'. *Work, Employment & Society*, *7*(2), 213-235.

Hasan, R., & Jandoc, K. (2009). Quality of Jobs in the Philippines: Comparing Self-employment with Wage Employment. *ADB Economics Working Paper,* (148).

Heckman, J. (1979). Sample Selection Bias as a Specification Error. *Econometrica*, *47*(1), 153.

Hyclak, T., Johnes, G., & Thornton, R. (2005). *Fundamentals of Labor Economics*. Houghton Mifflin Company.

ILO (2008). *World of Work Report 2008: Income Inequalities in the Age of Financial Globalization, International Institute for Labour Studies.* International Institute for Labour Studies.

ILO (2012). *Decent Work Country Profile: Cambodia*. Geneva. Retrieved from http://www.ilo.org/wcmsp5/groups/public/---dgreports/---integration/documents/publication/wcms_185267.pdf

ILO (2013). *Meeting the challenge of precarious work: A workers' agenda*. Retrieved from http://www.ilo.org/wcmsp5/groups/public/---ed_dialogue/---actrav/documents/publication/wcms_216282.pdf

ILO (2015a). Non-standard forms of employment: Report for discussion at the meeting of experts on non-standard forms of employment. *ILO Conditions of Work and Equality Department MENSFE/2015*.

ILO (2015b). Pakistan Decent Work Country Programme 2010-2015. Retrieved from http://www.ilo.org/public/english/bureau/program/dwcp/download/pakistan2010-15.pdf

ILO (2015c). Philippine employment trends 2015- Decent jobs crucial for inclusive growth. Retrieved from http://www.ilo.org/manila/publications/WCMS_362751/lang--en/index.htm

ILO (2015d). Nothing more permanent than temporary? Understanding fixed-term contract. Policy Brief N° 6. Geneva. Retrieved from http://www.ilo.org/wcmsp5/groups/public/---ed_protect/---protrav/---travail/documents/publication/wcms_357403.pdf

ITUC (2014). *Precarious work in the Asia Pacific Region. A 10 country study by The International Trade Union Confederation (ITUC) and ITUC Asia-Pacific*. Retrieved from http://www.ituc-csi.org/IMG/pdf/report_2014__precarious_work_lr.pdf

Jahn, E. & Pozzoli, D. (2013). The Pay Gap of Temporary Agency Workers — Does the Temp Sector Experience Pay off?. *Labour Economics*, *24*, 48-57.

Kalleberg, A., Reynolds, J., & Marsden, P. (2003). Externalizing Employment: Flexible Staffing Arrangements in US Organizations. *Social Science Research*, *32*(4), 525-552.

Kemal, A. & Mahmood, Z. (1998). The urban informal sector of Pakistan: Some stylized facts. *Pakistan Institute of Development Economics Research report No. 161*.

Kim, H. & Lee, D. (2014). The Relationship between the Internal Labour Market and Transitions from Temporary to Permanent Employment in Korea. *Relations Industrielles*, *69*(3), 597-620.

Ko, J. (2003). Contingent and Internal Employment Systems: Substitutes or Complements?. *Journal of Labor Research*, *24*(3), 473-490.

Lee, B.-H. & Lee, S. (2007). Minding the gaps: Non-regular employment and labour market segmentation in the Republic of Korea. *ILO Conditions of Work and Employment Series 19*.

Lee, S. & Eyraud, F. (2008). *Globalization, Flexibilization and Working Conditions in Asia and the Pacific*. Chandos Publishing.

Machado, J. & Mata, J. (2005). Counterfactual Decomposition of Changes in Wage Distributions Using Quantile Regression. *Journal of Applied Econometrics*, *20*(4), 445-465.

McGinnity, F., Mertens, A., & Gundert, S. (2005). A Bad Start? Fixed-term Contracts and the Transition from Education to Work in West Germany. *European Sociological Review*, *21*(4), 359-374.

Mertens, A., Gash, V., & McGinnity, F. (2007). The Cost of Flexibility at the Margin. Comparing the Wage Penalty for Fixed-term Contracts in Germany and Spain using Quantile Regression. *Labour*, *21*(4-5), 637-666.

Nawata, K., & Ii, M. (2004). Estimation of the Labor Participation and Wage Equation Model of Japanese Married Women by the Simultaneous Maximum Likelihood Method. *Journal of the Japanese and International Economies*, *18*(3), 301-315.

Nesheim, T., Olsen, K., & Kalleberg, A. (2007). Externalizing the Core: Firms' Use of Employment Intermediaries in the Information and Communication Technology Industries. *Human Resource Management*, *46*(2), 247-264.

Neuman, S., & Oaxaca, R. (2004). Wage Decompositions with Selectivity-Corrected Wage Equations: A Methodological Note. *The Journal Of Economic Inequality*, *2*(1), 3-10.

Nguyen, H., Nordman, C., & Roubaud, F. (2013). Who Suffers the Penalty?: A Panel Data Analysis of Earnings Gaps in Vietnam. *Journal Of Development Studies*, *49*(12), 1694-1710.

Oaxaca, R. (1973). Male-female Wage Differentials in Urban Labor Markets. *International Economic Review, 14*(3), 693–709.

OECD (2014). *OECD Employment Outlook 2014*. OECD Publishing.

Pacheco, G., & Cochrane, B. (2015). Decomposing the Temporary-permanent Wage Gap in New Zealand. *AUT Department of Economics Working Paper Series 2015/07*.

Puhani, P. (2000). The Heckman Correction for Sample Selection and Its Critique. *Journal Of Economic Surveys*, *14*(1), 53-68.

Rani, U. (2008). Impact of Changing Work Patterns on Income Inequality. *International Institute for Labour Studies - Discussion Paper Series DP/193/2008*.

Rebitzer, J. & Taylor, L. (1991). A Model of Dual Labor Markets When Product Demand Is Uncertain. *The Quarterly Journal Of Economics*, *106*(4), 1373-1383.

Reich, M., Gordon, D., & Edwards, R. (1973). A Theory of Labor Market Segmentation. *American Economic Review, 63*(2), 359–365.

Rosen, S. (1986). The theory of equalizing differences. *Handbook of labor economics*, *1*, 641-692.

Salvatori, A. (2009). What do unions do to temporary employment. *IZA Discussion Paper No. 4554*.

Serrano, M., Marasigan, M., & Pupos, V. (2014). *Between Flexibility and Security: The Rise of Non-Standard Employment in Selected ASEAN Countries*. ASEAN Services Employees Trade Unions Council (ASETUC).

Tucker, D. (2002). *'Precarious' Non-Standard Employment - A Review of the Literature*. Labour Market Policy Group, Department of Labour.

USAID (2006). *Cambodia Garment Industry Workforce Assessment: Identifying Skills Needs and Sources of Supply*. Retrieved from https://cstn.files.wordpress.com/2009/12/cambodian-garment-industry-workforce-assessment-june06-en.pdf.

Appendix

A. Data description

Data used in the study obtained from national LFS provided by ILO. Further details for individual countries are as follows:

Cambodia

The Labor Force and Child Labor Survey (LF-CLS) was conducted by the National Institute of Statistics, Ministry of Planning in 2011-2012. This survey provides data on labor force, employment, unemployment, and other data for monitoring and assessing progress on decent work of population aged 15 years and above.

Indonesia

The survey's name abbreviates SAKERNAS. Since 2005, SAKERNAS has been conducted biennially (in February and in August) to reflect seasonal labor force variations. The dataset in the current study come from the 2007 round conducted in August.

Pakistan

Since the LFS 2005-2006, the questionnaire has been being articulated for undertaking quarterly representative results reflecting not only labor force characteristics but also important related attributes of literacy, migration, occupational safety, etc. at national level. The data used in this study is extracted from 2008-2009 wave.

The Philippines

Data come from the 2009 round of the annual LFS in the Philippines.

Vietnam

The dataset is obtained from LFS 2007 conducted by the Vietnam General Statistics Office (GSO). The 2007 LFS marked the transfer of the LFS from the Ministry of Labor, Invalids and Social Affairs to the GSO to build up an improved scheme providing more complete labor market indicators meeting international standards.

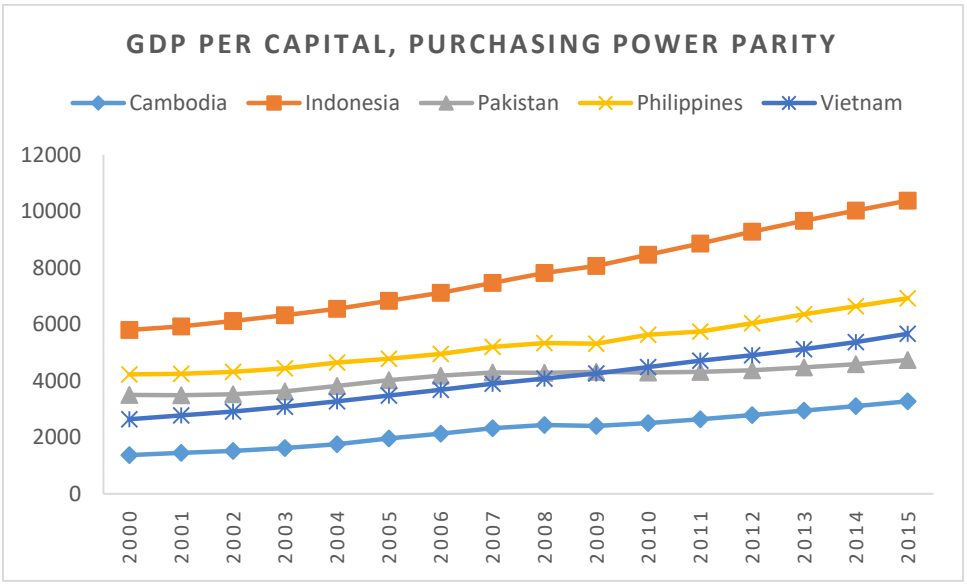
Table 14. Types of Non-standard Workers and Terms Used Referring to Non-standard Employment in the Formal Sector

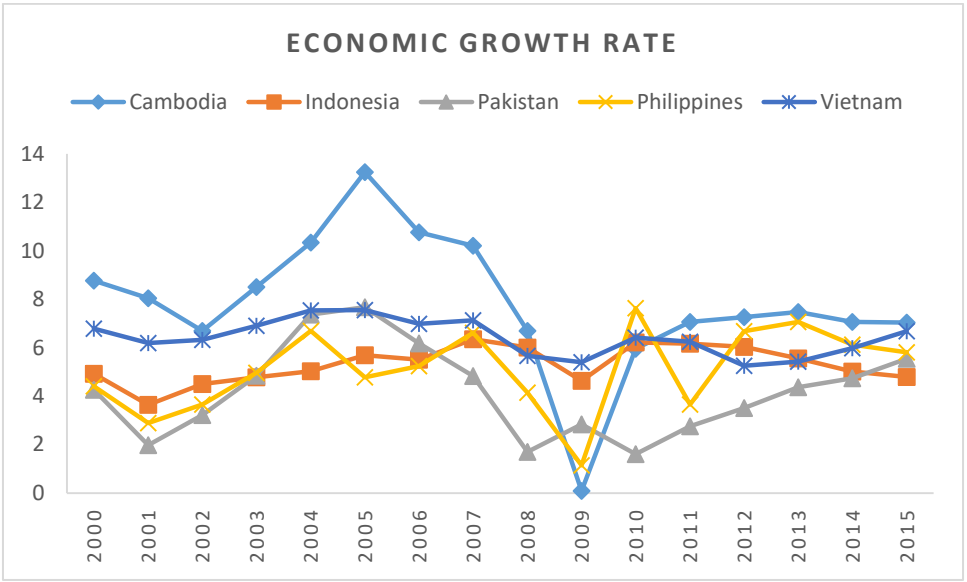
Country	Types of non-standard workers/ Terms used	Non-standard employment involving triangular employment relations
Indonesia	Short/fixed-term contract workers Casual workers Outsourced workers Apprentices Part-time workers Piece rate workers On-call workers	Outsourced work
Malaysia	Fixed-term contract Contract of service Contract for service Casual paid employee Contract/piece paid worker	Outsourced work, contractor or subcontractor for labor, agency-hired work
Philippines	Short-term contract work (i.e., casual, contractual/project-based, seasonal, probationary, apprenticeship/learnership) Outsourcing or subcontracting of work, agency hired and third-party managed work	Agency-hired work, manpower supply, service or labor contracting
Singapore	Term contracts Casual/on-call employment Part-time work	Outsourced workers
Thailand	Hire of service employment contract Hire of work contract Informal employment Dispatched, outsourced and agency employment	Contracting/subcontracting, labor contracting
Vietnam	Seasonal work Labor dispatch or employee subleasing	Labor dispatch or employee subleasing

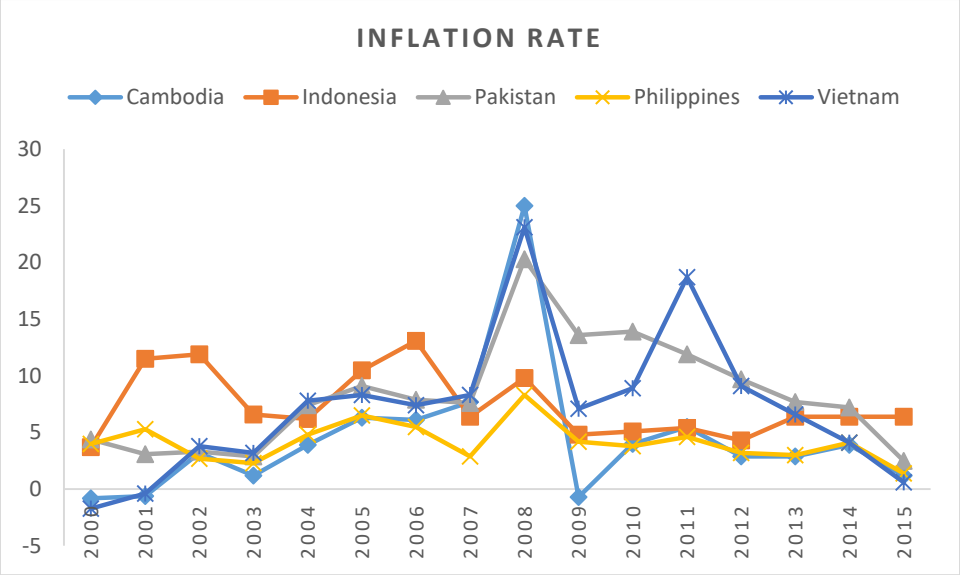
Source: Serrano et al., 2014.

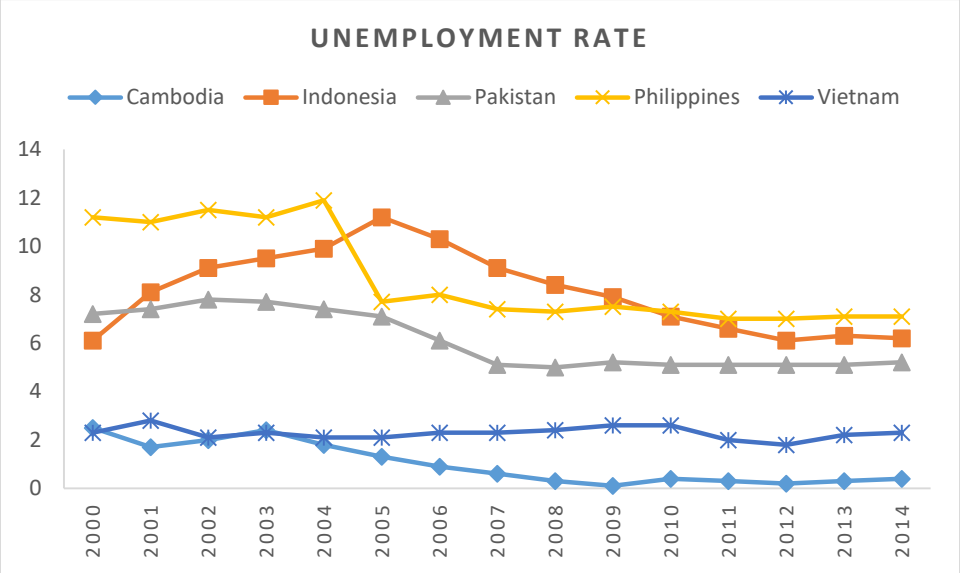
B. Socio-economic conditions in five selected Asian countries

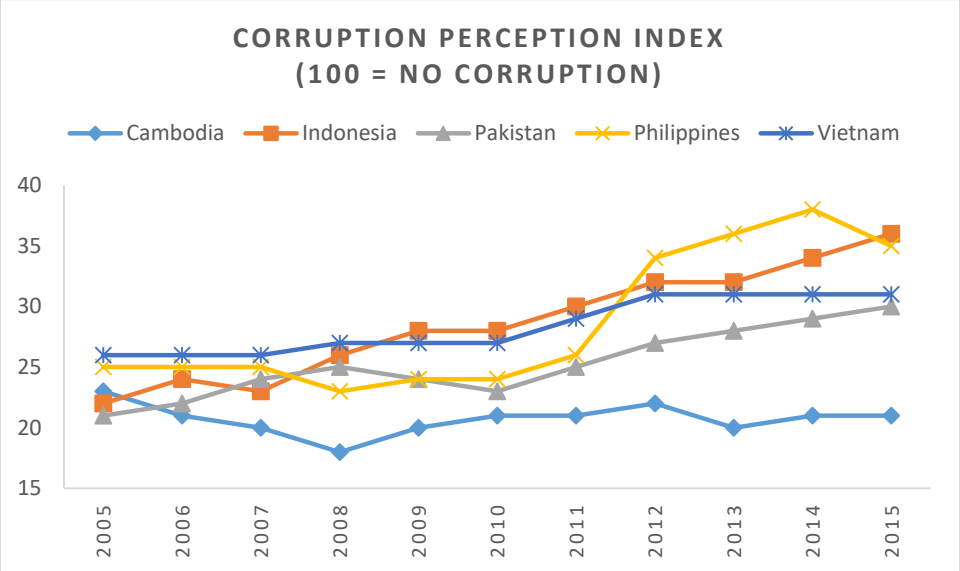
Figure 3. Socio-economic Conditions in Five Selected Asian Countries

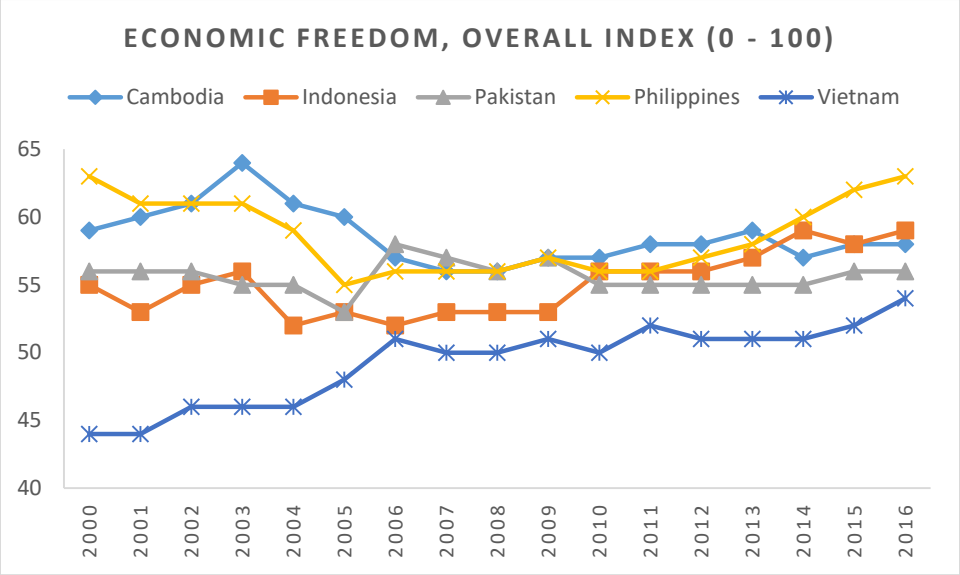


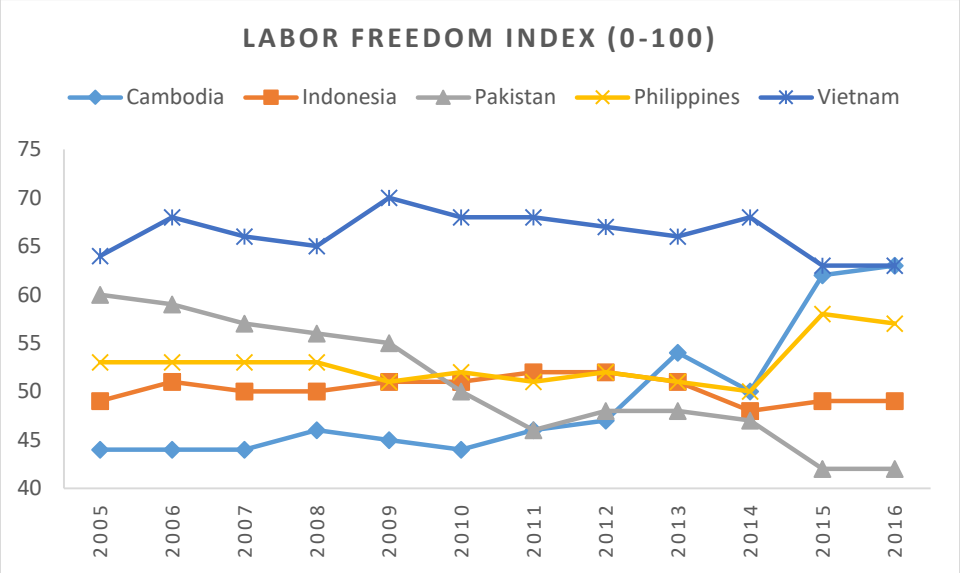


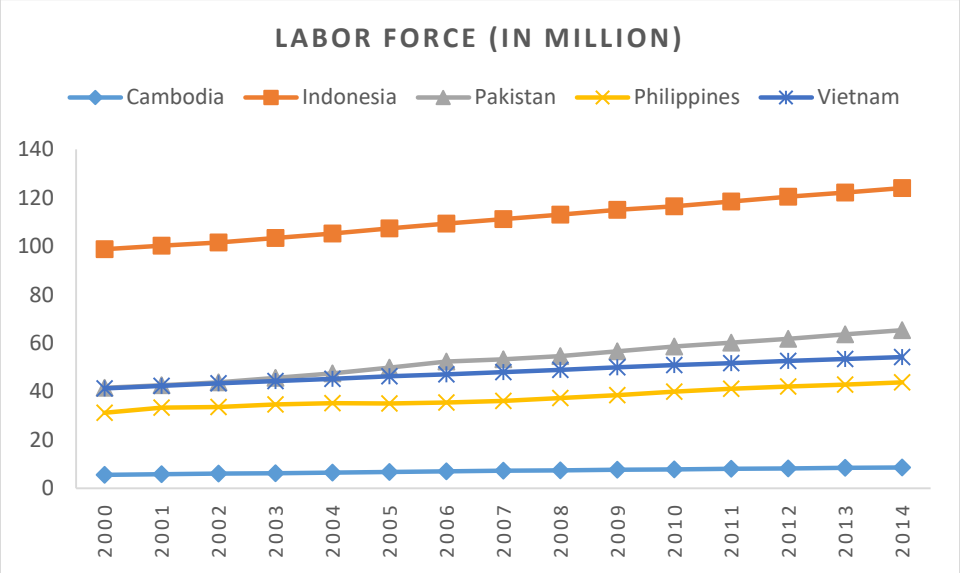

INFLATION RATE
Cambodia
Indonesia
Pakistan
Philippines
Vietnam
30
25
20
15
10
5
0
-5
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015

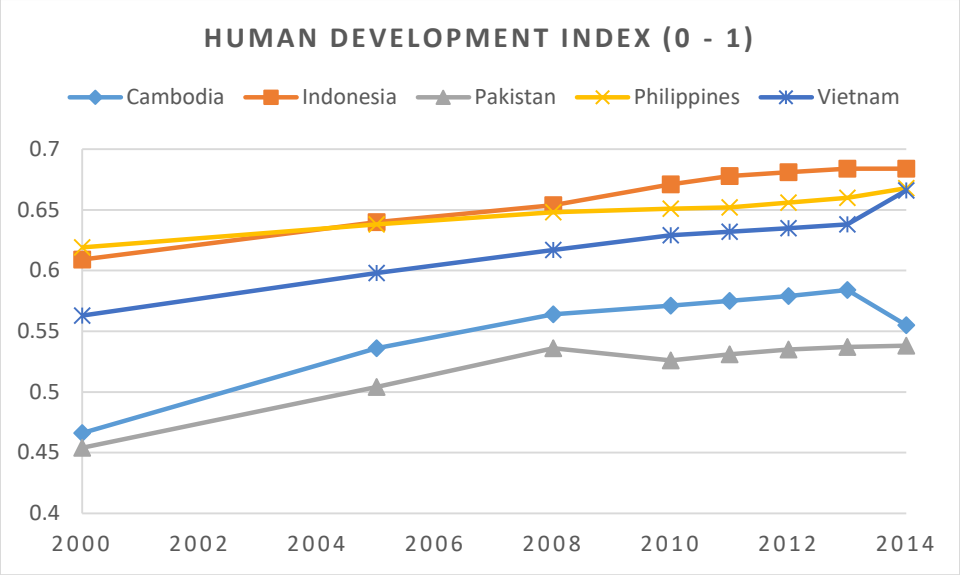

UNEMPLOYMENT RATE
Cambodia
Indonesia
Pakistan
Philippines
Vietnam
14
12
10
8
6
4
2
0
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014

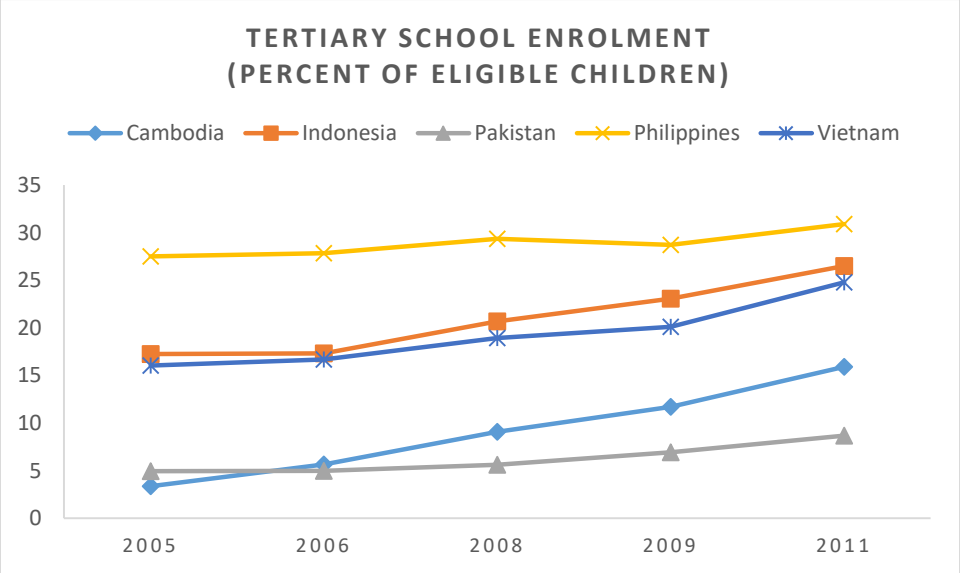

CORRUPTION PERCEPTION INDEX
(100 = NO CORRUPTION)
Cambodia
Indonesia
Pakistan
Philippines
Vietnam
40
35
30
25
20
15
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015


ECONOMIC FREEDOM, OVERALL INDEX (0 - 100)
Cambodia
Indonesia
Pakistan
Philippines
Vietnam
65
60
55
50
45
40
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016


LABOR FREEDOM INDEX (0-100)
Cambodia
Indonesia
Pakistan
Philippines
Vietnam
75
70
65
60
55
50
45
40
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016


LABOR FORCE (IN MILLION)
Cambodia
Indonesia
Pakistan
Philippines
Vietnam
140
120
100
80
60
40
20
0
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014





Data source: theglobaleconomy.com, the World Bank. Figures constructed by authors.

Table of Contents

General Introduction

1. General context 10
   1.1. Asia: Yesterday's glory, today's reemergence and the question of inclusive growth 10
   1.2. Vietnam: An ugly duck turns into a swan? 12
2. Chapters' summary and contextualization 14
   2.1. Chapter 1: Heterogeneity of the informal sector in Vietnam: A *'quali-quanti'* approach 14
   2.2. Chapter 2: Returns to higher education in Vietnam: Who benefit the most? 20
   2.3. Chapter 3: Wage differentials between temporary and permanent workers in Asian developing countries 24

Chapter 1: Heterogeneity of the informal sector in Vietnam: A *'quali-quanti'* approach

Abstract 31
1. Introduction 32
2. Literature Review 33
3. Data and methodology 35
   3.1. Data 35
   3.2. Methodology 38
4. Findings 41
   4.1. Descriptive statistics 41
      HBs and informal HBs: Who are they? 41
      First insights from the textual data of HB&IS 2009 49
   4.2. Correspondence Analysis on the open-ended questions 53
      Advantages 54
      Disadvantages 56
      Combined corpus of advantages and disadvantages 65
   4.3. Cluster analysis 66
   4.4. Econometric models: How textual data can explain HB's economic performance? 70
      The relationship between 'entrepreneurial skill' and economic performance of HBs 70
      Factorial axes as a measure of entrepreneurial skill 82
   4.5. Robustness check 83
5. Conclusion 87
Appendix 89
References 94

Chapter 2: Returns to higher education in Vietnam: Who benefit the most?

Abstract 99
1. Introduction 100
2. Literature review 101
3. Data and descriptive statistics 106
3.1. Data 106
3.2. Descriptive Statistics 107
4. Methodology 112
4.1. Potential bias of OLS estimation on returns to education 112
4.2. Empirical Strategies 113
4.2.1. Ordinary Least Square 113
4.2.2. Instrumental Variable Estimation 115
4.2.3. Heterogeneous returns to higher education 116
5. Average returns to higher education: OLS and IV estimates 117
6. Heterogeneity of returns to higher education in Vietnam 120
6.1. OLS estimates 120
6.2. Quantile Regressions 124
6.3. Heterogeneous Treatment Effects 127
7. Robustness checks 133
8. Discussion and conclusion 137
Appendix 140
References 149

Chapter 3: Wage differentials between temporary and permanent workers in Asian developing countries

Abstract 153
1. Introduction 154
2. Literature review 155
2.1. Theoretical background 155
Why workers and firms seek NSFE? 155
Temporary workers: wage penalty or wage premium? 156
Sticky floor versus *glass ceiling* effects 158
2.2. Empirical evidence on temporary-permanent workers' wage gap 158
3. Country context, data and methodology 159
3.1. A brief introduction of the economy and labor market context in selected Asian countries 159

3.2. Data and descriptive statistics 162
Data and variables 162
Statistical description 164
3.3. Estimation strategies 169
4. Econometric analysis of wage differentials 171
4.1. Estimates at mean: OLS and Heckman models 171
4.2. Quantile estimation analysis 174
4.3. How we define 'temporary employment' matters 190
4.4. Discussion and interpretation of the estimation results 193
Indonesia 195
Vietnam and Pakistan 195
The Philippines 196
Cambodia 197
5. Concluding remarks 201
References 204
Appendix 209
A. Data description 209
B. Socio-economic conditions in five selected Asian countries 211
Centre for Analysis and Forecasting 222
Institut de Recherche pour le Développement 222
Vietnam Academy Of Social Sciences 222

List of Tables

General Introduction

Table 1. Educational Attainment of the Population Aged 20 to 24 Years by Household Wealth in Viet Nam, 2011. 21

Chapter 1: Heterogeneity of the informal sector in Vietnam: A *'quali-quanti'* approach

Table 1. Main characteristics of HBs and informal sector in HCMC (in percentage) 41
Table 2. Numerical Coding of Texts: A Summary 49
Table 3. Number of responses to the close-ended question versus to the open-ended question concerning the problems and disadvantages of informal HBs 50
Table 4. Most Frequent Words Used by Informal HBs 51
Table 5. The Main Difficulties Faced by Informal HBs 52
Table 6. The First Five Eigenvalues of Correspondence Analysis 54
Table 7. Selective Points with Extreme Contributions to the First Two Principal Axes: Correspondence Analysis on "Disadvantages" of Informal HBs 55
Table 8. Selective Points with Extreme Contributions on the First Two Principal Axes: Correspondence Analysis on "Disadvantages" of Informal HBs 56
Table 9. Selective Points with Extreme Contributions on the First Two Principal Axes: Correspondence Analysis on "Disadvantages" of HBs 58
Table 10. Selective Points with Extreme Contributions on the First Two Principal Axes: correspondence analysis on "Advantages" and "Disadvantages" of Informal HBs 65
Table 11. Positions of Clusters in the Dimensional Space –Disadvantages of Informal HBs 67
Table 12. Categorical Characteristics of the Clusters – Cluster Analysis on Disadvantages of Informal HBs 68
Table 13. Determinants of Economic Performance of HBs 73
Table 14. Determinants of Total Size of HBs 76
Table 15. Determinants of Investment in Equipment of HBs 79
Table 16. 'Entrepreneurial capital' and loan usage behavior 82
Table 17. Selective Points with Extreme Contributions on the First Two Principal Axes: Correspondence Analysis on "Disadvantages" and Motivations of Informal HBs 84
Table 18. Categorical Variables Used in Correspondence Analysis 89
Table 19. Mean Tests for the Non-responses to Open-ended Questions 92
Table 20. Positions of Clusters in the Dimensional Space – Advantages of Informal HBs 92
Table 21. Characteristic Responses of the Clusters – Cluster Analysis on Advantages of Informal HBs 92
Table 22. Categorical Characteristics of the Clusters – Cluster Analysis on Advantages of Informal HBs 93

Chapter 2: Returns to higher education in Vietnam: Who benefit the most?

Table 1. Higher Education Coverage in Vietnam (unrestricted sample) 108
Table 2. Higher Education Coverage in Vietnam (restricted sample) 109
Table 3. Descriptive Comparison between Non-Treated and Treated Groups 110
Table 4. Average Hourly Income in Different Groups 111
Table 5. Returns to Higher Education in Vietnam. Dependent Variable: Log Hourly Earnings 119
Table 6. OLS Estimates of Returns to Education in Vietnam for Different Sub-populations, without Family Background Controls 121

Table 7. OLS Estimates of Returns to Education in Vietnam for Different Sub-populations, with Family Background Controls 122
Table 8. Returns to Higher Education for Females in Difference Ages 123
Table 9. Returns to Higher Education in Vietnam - Quantile Regressions 126
Table 10. Who Are More Likely to Go to College? Probit Estimations 128
Table 11. Returns to Higher Education by Propensity Strata 129
Table 12. Returns to Higher Education in Vietnam – Robustness Checks 134
Table 13. Sub-sample with Available Information on Family Background: Is There Sample Selection? 135
Table 14. Are Minor Ethnic High School Graduates Truly Exceptional? 136
Table 15. High School Non-graduates versus Graduates: Kinh People Only 136
Table 16. Literature on Returns to Education in Vietnam 140
Table 17. Vietnam's Labor Force Survey in Different Rounds: A Comparative View 145

Chapter 3: Wage differentials between temporary and permanent workers in Asian developing countries

Table 1. Definition of Temporary Employment in National Labor Force Surveys 166
Table 2. Temporary Employees in Selected Asian Countries: Statistical Description 167
Table 3. Temporary Workers' Wage Gap in Selected Asian Countries – Benchmark OLS Model 172
Table 4. Temporary-Permanent Workers' Wage Gap: OLS and Quantile Regression 176
Table 5. Temporary-Permanent Workers' Wage Gap: Heckman Correction for Sample Selection 177
Table 6. Temporary-Permanent Workers' Wage Gap: Extended OLS and Quantile Regression - Cambodia. 179
Table 7. Temporary-Permanent Workers' Wage Gap: Extended OLS and Quantile Regression – Indonesia 181
Table 8. Temporary-Permanent Workers' Wage Gap: Extended OLS and Quantile Regression – Pakistan 183
Table 9. Temporary-Permanent Workers' Wage Gap: Extended OLS and Quantile Regression – The Philippines 185
Table 10. Temporary-Permanent Workers' Wage Gap: Extended OLS and Quantile Regression – Vietnam 187
Table 11. Temporary-Permanent Workers' Wage Gap: OLS and Quantile Regression – Modified Definitions of Temporary Employment 192
Table 12. Macroeconomic Conditions in Selected Asian Countries: A Comparative View 194
Table 13. Temporary-Permanent Workers' Wage Gap: OLS and Quantile Regression, Extended Models by Gender 199
Table 14. Types of Non-standard Workers and Terms Used Referring to Non-standard Employment in the Formal Sector 210

List of Figures

General Introduction

Figure 1. Intensity of the Informal Sector by Province, 2007-2014 16
Figure 2. Main Problems Encountered by HBs, by Institutional Sector (Percent of HBs) 17
Figure 3. Motivation of Setting Up an HB (percent of HB) 18
Figure 4. Satisfaction of Working in the HB and Informal Sector 19
Figure 5. R&D Expenditure as a Percentage of GDP and Researchers in Full-time Equivalents (FTE) per 1 Million Inhabitants, 2011 or Most Recent Year Available 22
Figure 6. Job Quality and Job Quantity Outcomes by Type of Employment: Cross-country Averages, 2010. 25
Figure 7. Incidence of Temporary Workers in Selected Asian Countries, as Percentage of Wage Workers, by Level of Education 26

Chapter 1: Heterogeneity of the informal sector in Vietnam: A *'quali-quanti'* approach

Figure 1. Distribution of HBs According to Economic Performance (in Thousand Vietnam Dongs) 46
Figure 2. The Space of Informal HB Characteristics (First Factorial Plane of Correspondence Analysis on Advantages) 60
Figure 3. The Space of Disadvantages Put Forward by Informal HBs (First Factorial Plane of Correspondence Analysis) 61
Figure 4. The Space of Informal HB Characteristics (First Factorial Plane of Correspondence Analysis on Disadvantages) 62
Figure 5. The Space of Disadvantages Put Forward by HBs (First Factorial Plane of Correspondence Analysis) 63
Figure 6. The Space of HB Characteristics (First Factorial Plane of Correspondence Analysis on Disadvantages) 64
Figure 7. Cluster View of Informal HBs on the First Factorial Plane of Correspondence Analysis on Disadvantages 69
Figure 8. The Space of Informal HB Characteristics (First Factorial Plane of Correspondence Analysis on Disadvantages and Motivation) 86

Chapter 2: Returns to higher education in Vietnam: Who benefit the most?

Figure 1. Returns to Higher Education in Vietnam along the Earnings Distribution 124
Figure 2. Quantile Regression without Family Background Controls: Public versus Private Sectors 127
Figure 3. Returns to Higher Education in Vietnam across Propensity Score Strata 131

Chapter 3: Wage differentials between temporary and permanent workers in Asian developing countries

Figure 1. Temporary-Permanent Workers' Wage Differentials in the Selected Asian Countries, at Mean and Different Quantiles. Extended Set of Explanatory Variables. 189
Figure 2. Temporary-Permanent Workers' Wage Differentials by Gender in the Selected Asian Countries, at Mean and Different Quantiles. Extended Set of Explanatory Variables 200
Figure 3. Socio-economic Conditions in Five Selected Asian Countries 211

List of Acronyms

ASEAN	Association of Southeast Asian Nations
ATE	Average Treatment Effect
CAF	Centre for Analysis and Forecasting
DIAL	Développement, Institutions & Mondialisation
FTC	Fixed-term Contract
FTE	Full-time Equivalents
GDP	Gross Domestic Product
GSO	General Statistics Office
HB	Household Business
HB&IS	Household Business and Informal Sector
HCMC	Ho Chi Minh City
ILO	International Labor Organization
IRD	Institut de Recherche pour le Développement
ITUC-AP	International Trade Union Confederation – Asia Pacific
IV	Instrumental Variable
LFS	Labor Force Survey
NSFE	Non-standard Forms of Employment
OECD	Organization for Economic Co-operation and Development
OLS	Ordinary Least Square
PCA	Principal Component Analysis
PU	Production Unit
R&D	Research and Development
SPOFS	Social Protection Operational Framework and Strategy
SUTVA	Violation of Stable Unit Treatment Value Assumption
USAID	United States Agency for International Development
VASS	Vietnam Academy Of Social Sciences
VGCL	Vietnam General Confederation of Labor
VHLSS	Vietnam Household Living Standards Survey
VLSS	Vietnam Living Standards Survey

Résumé

Cette thèse de doctorat porte sur la croissance «exclusive» et la vulnérabilité qui caractérisent le développement du Vietnam et plus largement de l'Asie aujourd'hui. Les chapitres traitent trois aspects importants de la vulnérabilité et du développement inclusif, à savoir : l'informalité (chapitre 1), le dilemme de l'éducation (chapitre 2) et l'emploi atypique (chapitre 3). Les contributions de ce travail reposent sur différents facteurs : la nouveauté et la pertinence des sujets de recherche ; le large éventail de données utilisées, quantitatives et qualitatives, y compris les Enquêtes sur les entreprises individuelles et sur le Secteur Informel au Vietnam, et les Enquêtes Nationales sur l'Emploi de divers pays d'Asie ; et enfin l'originalité de la méthodologie. Le chapitre 1 étudie l'hétérogénéité du secteur informel au Vietnam, en se basant sur une approche quali-quantitative unique. Le chapitre 2 se concentre sur la variation des rendements de l'enseignement supérieur dans la population vietnamienne en recourant à différents modèles d'estimation. Le chapitre 3 est la première étude qui étudie systématiquement les écarts de salaire induits par le statut de travail temporaire dans les pays en développement d'Asie. Dans l'ensemble, toute la thèse implique que le capital humain, l'emploi et le revenu sont des facettes interdépendantes du bien-être individuel et que certains phénomènes de développement doivent être analysés dans leur hétérogénéité.

Mots Clés

Vulnerabilité, Développement inclusif, Informalité, Rendements de l'éducation, Emploi atypique, Asie en développement, Vietnam

Abstract

This PhD dissertation is dedicated to the issue of 'exclusivist' growth and vulnerability characterizing Vietnam as well as developing Asia today. The chapters address three important aspects of vulnerability and inclusive development, namely: Informality (chapter 1), Education dilemma (chapter 2) and Non-standard employment (chapter 3). The contribution of this work lies in the novelty and relevance of research topics; the wide range of data used, both quantitative and qualitative, including Household Business and Informal Sector Surveys in Vietnam, and national Labor Force Surveys of various countries in Asia; as well as the originality of methodology. Chapter 1 investigates the heterogeneity of the informal sector in Vietnam, based on a unique quali-quanti approach. Chapter 2 focuses on the variation of the returns to higher education across the Vietnamese population with different estimation models. Chapter 3 is the first study that systematically examines the wage differentials induced by temporary job status in Asian developing countries. Overall, the whole thesis implies that human capital, employment, and income are interrelated facets of individual well-being, and that some development phenomena should be analyzed in their heterogeneity.

Keywords

Vulnerability, Inclusive development, Informality, Returns to education, Non-standard employment, Developing Asia, Vietnam