An analysis of labour market returns to education in Vietnam: evidence from the national labour force survey 2012

Nguyen Dung Tien

TURIN SCHOOL of DEVELOPMENT

ITC ILO
International Training Centre

International Training Centre of the International Labour Organization
Viale Maestre del Lavoro, 10
10127, Turin, Italy
www.itcilo.org

The Working Paper of the Turin School of Development is part of a wider effort by the International Training Centre of the ILO to contribute to the development policy dialogue through a wide dissemination of research findings of its staff, postgraduate course participants, partner institutions and external collaborators. Papers are selected for publication on the basis of originality, technical soundness and policy relevance.

This paper was originally written by Nguyen Dung Tien under the supervision of Luigi Benfratello in partial fulfilment of the requirement of the Master in Applied Labour Economics for Development offered by the Turin School of Development of the International Training Centre of the ILO in partnership with the University of Turin and the Institut d'Etudes Politiques de Paris (Sciences Po). More information on the Masters Programme is available at http://maled.itcilo.org/

Publication in the Working Paper of the Turin School of Development does not preclude the authors from subsequently publishing the papers in peer-reviewed journals, books and other venues.

Suggested citation for this paper

Tien, N. D. L. 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.

AN ANALYSIS OF LABOUR MARKET RETURNS TO EDUCATION IN VIETNAM: EVIDENCE FROM THE NATIONAL LABOUR FORCE SURVEY 2012

Nguyen Dung Tien

Corresponding author: E-mail: nguyendungtien@gmail.com

Blank page

Abstract

The importance of education to a country's development strategy has been widely acknowledged. In Vietnam, education is considered a national policy of top priority. Because of this, in this country the analysis of the returns to education is a popular theme. However, most studies focus on wage premiums for education using Vietnam Household Living Standards Surveys from years prior to 2008. This paper uses the National Labour Force Survey of 2012 to provide an updated and detailed examination of the returns to education in the labour market, taking into account the heterogeneity of a developing country context.

Blank page

Contents

Blank page

1. Introduction

Over the past 20 years, Vietnam's development has generally been considered a success story. Political and economic reforms launched in 1986 have transformed Vietnam from one of the poorest countries in the world, with per capita income below $100 per annum, to a lower-middle income country, with per capita income of $1,130 per annum by the end of 2010. The equity of its development has been better than in most other countries in similar situations (World Bank, 2014).

In a changing reality like Vietnam, education and training play an important role in the country's development strategy. Education is not only a pathway to poverty reduction, but it also equips the workforce with the necessary skills to transform an agriculture-based economy into an industrialized economy. For this reason, the Government of Vietnam considers national education policy to be a top priority. Important achievements have been made in the programme of universalization of primary and lower secondary education. Moreover, students have access to loans for tertiary education. However, despite the achievements in improving access to education for the masses, many remain sceptical about its low quality and relevance (MOET, 2014).

Measuring the extent to which investment in education leads to gains in terms of labour market performance is important to informing public policy. Returns to education in Vietnam have been discussed extensively in the literature. However, most studies focus on the wage premiums attributable to education using Vietnam Household Living Standards Surveys conducted prior to 2008. For a developing country in a transition period such as Vietnam, it has been argued that the rate of return to education tends to increase (Flabbi et al., 2008). It is therefore essential to re-examine the linkage between education and labour market performance employing updated data.

Using the National Labour Force Survey of 2012, this paper aims to provide a more up-to-date and detailed examination of the impact of education on labour market outcomes, taking into account the heterogeneity of a developing country context. The National Labour Force Survey has been conducted in Vietnam by the General Statistical Office (GSO) on a quarterly basis since 2011 with technical support from the International Labour Organization. The main objective of the survey is to collect data on key indicators related to Vietnam's labour market, including those employed, those unemployed, and those not in the labour force, and to provide descriptive and explanatory data on each category. The panel data provided by this high-frequency survey offers a very good opportunity to re-visit the issue of measurement of private returns to education in Vietnam. Since this is the first study to use the National Labour Force Survey, this paper may also serve as a reference for future studies analysing trends in returns to education in Vietnam, using the same data.

1.1. Structure of the paper

Chapter 2 sets out the theoretical background of the analysis. It first explains the human capital theory, the basic Mincer model for estimating private returns to education, before providing a review of the literature on improvements to the Mincer equations. Finally, the studies undertaken on returns to education in Vietnam will be reported. The chapter concludes by highlighting important issues in the estimation of private returns to education in Vietnam.

Chapter 3 introduces the data from the Vietnamese Labour Force Survey 2012 and addresses fundamental methodology issues in the estimation of private returns to education in this country, given the available data. A regression model will be proposed.

Chapter 4 discusses the results of the analysis of the labour market outcomes attributable to education. Descriptive statistics are presented, followed by a more detailed examination of how investments in education pay off in terms of labour market earnings, subdivided according to economic sector, ownership, and gender.

Chapter 5 summarizes key findings and presents the study's policy recommendations.

1.2. Scope of the study

This study attempts to make use of all variables observed in the Labour Force Survey to account for the benefits to labour market earnings attributable to education. Unobserved issues are not discussed, for example the asymmetric information between worker and employer (employers generally cannot observe the net potential earnings of workers), the role of trade unions, efficiency wages, and so on.

2. A review of the literature

This chapter is dedicated to an in-depth review of the theoretical framework of this study. Human capital theory and the Mincer model will be described in order to define the economic and econometric approaches used in this type of analysis. Previous findings on returns to education in Vietnam will be discussed at the end of the chapter.

2.1. Human capital theory

The theoretical basis for this study is human capital theory. Human capital consists of the intangible set of skills and knowledge that a worker possesses, including those that are innate and those acquired through schooling, non-schooling investments, pre-labour market influences and labour market experience (Becker, 2008). Investments in education and training help to improve these intangible skills and, in turn, increase a worker's productivity. Education and training also equip workers with the skills required for the labour market, enabling them to be promoted in their employment or find a better job, and increase their income. Education and training is therefore considered a short-term investment for long-term development and one of the most important variables in determining the prosperity of individuals and the wealth of nations.

Empirical research on the causal effects of education on labour market outcomes has been conducted in many countries around the world. Data across countries and different time periods consistently proves that better-educated people earn higher wages and are less likely to be unemployed or occupy informal jobs than their less educated counterparts (Card, 1999).

Human capital theory explains why investment in education leads to better labour market outcomes. However, variations in the observed income of workers in the labour market are attributable to many other factors. In a developing country's labour market, a significant percentage of the workforce is self-employed or works in the informal economy. Education equips workers with knowledge and skills but their income level also depends, for example, on network capital (networks of friends, relatives, etc). In the formal sector, employers cannot precisely estimate the abilities or potential that an employee will develop in the future, or they may be constrained by financial difficulties at the time of recruitment. A number of such heterogeneity issues will be discussed in the next section.

2.2. Basic Mincerian model

The rate of return to education is important both for individual decision-making related to education and also for governments and policy-makers allocating scarce resources. Over the last 30 years a thorough body of literature has examined this linkage between education and earnings, almost all of which is based on the model developed by Jacob Mincer (1974).

In Mincer's original model, the variation in individuals' wages is explained by variations in years spent in education and years of labour market experience, with a linear relationship between schooling and wages.

$$Y = \beta_0 + \beta_1 X + \beta_2 Z + \beta_3 Z^2 + \mu \qquad (1)$$

Where: Y = log of workers' hourly wage

X = years of education

Z = years of experience

β_1 estimates the rate of return to schooling, the coefficient of interest. β_2 and β_3 estimate the returns to experience (reflecting the concavity of returns to experience when β_3 is negative).

Traditionally, in order to face the problematic issue of heteroskedasticity, earnings are log-transformed. In fact, variability in income increases with the educational level: more educated workers have more job opportunities, which leads to higher earning variability in the higher levels of education.

There are many critiques of the basic Mincerian equation, and improvements have been suggested (Card, 1999). Nevertheless, the equation still offers a good starting point for analysing labour market returns to education. The accuracy of the model relies on a series of assumptions which are often violated. The following section summarizes some of these shortcomings in the estimation of rates of return to education by means of the static Mincerian model, namely (i) endogeneity, (ii) omitted variable bias, (iii) non-linearity in the relationship between wages and education, and (iv) sample selection bias.

2.3. Augmented Mincerian equation

Endogeneity due to unobserved ability

One of the most frequently discussed weaknesses of the basic Mincerian equation concerns the extent to which schooling is endogenous due to an "ability bias". Students have different unobserved abilities. Those with higher ability can stay longer in school, and workers with higher ability are likely to present themselves better at job interviews, and therefore have better labour market outcomes. Unobserved abilities affect both schooling decisions and income, creating a bias in the estimation of the causal relationships between years of schooling and income.

There are different ways of responding to this problem. One approach is to include an explanatory variable in the regression that can capture innate ability, such as an IQ test result or Armed Forces qualification tests (Harmon et al., 2003; Belzil and Hansen, 2002). However, these variables are not usually available.

Another commonly used approach is to use an instrumental variable (IV) that correlates closely with schooling but is not correlated with ability or wage level. These variables can be "distance to school", "spouse's education", "parents' education", "early age smoking" (Lall and Sakallariou, 2010; Card, 2001). However, according to part of the literature, the use of instrumental variables can lead to measurement error, weak instrument bias, and co-linearity with other socio-economic statuses. This fact can create additional biases in the estimate (Arcand et al., 2005; Card, 1999; Heckman and Urzua, 2009). Other studies argue that endogeneity is not a major issue. A more robust approach is to stick with OLS, while including an "ability"-related explanatory variable (Griliches, 1997; Angrist and Krueger, 1991). Endogeneity of schooling and approaches to the issue remain controversial (Heckman et al., 2003).

Omitted variable bias

A number of studies suggest that the basic Mincerian model overestimates returns to education owing to omitted variable biases other than unobserved ability (Lam and Schoeni, 1993; Card, 1999). When defining employees' earning levels, employers take into account other socio-economic factors that can determine the workers' potential in the future. For example, on-the-job training can affect an individual's ability to earn. In addition, an employer's decision on wages is constrained by the financial situation of the enterprise he runs. This suggests that employer-related variables, such as economic sector, ownership, or firm size, may have explanatory power in the regression model. Socio-economic factors can also improve the learning environment, or affect workers' wages through family connections to the job market. These factors are related to personal characteristics, such as gender, marital status and ethnicity. Including these variables, either as explanatory variables or controlling variables in (1) can improve the robustness of the result.

In this case, the extended Mincerian equation will be:

$$Y = \beta_0 + \beta_1 X + \beta_2 Z + \beta_3 Z^2 + \beta_4 T + \mu \tag{2}$$

Where:
Y = log of hourly wage

X = years of education

Z = years of experience

T = a vector of other explanatory and controlling variables

Non-linearity of education premium:

In the original model it is assumed that the log of hourly wage and years of schooling have a linear relationship. However, more and more evidence shows that the rate of return to schooling is convex (Mincer, 1997; Deschenes, 2001; Lemieux, 2003; Belzil, 2006). This is because different economic sectors demand different types of qualification. In some sectors there may be more demand for higher levels of education, leading to higher premiums for skilled workers. Another reason is the "sheepskin" effect (Mincer, 1997): employers value completed qualifications more than non-credentialed years. The rate of return for a student who has completed four years of bachelor studies and has a degree will be relatively higher than for another who has only completed three years and lacks a certificate.

The common approach to this concern is to use dummy variables for each level of education (e.g. primary, secondary, tertiary) to estimate the rate of return to each level. Another approach is to include the years of schooling squared in the model. The sign of the coefficient will show whether the education premium is concave or convex.

Sample selection bias

Another body of literature has proposed improvements to the Mincerian model on a sample selection bias, first developed by Heckman (1979). The intuition here is that we can only observe the income of those in employment, as data on those who are unemployed or unpaid will not be captured. The result of the model, therefore, only reflects the rate of return for a subset of the population, namely the population in employment. This creates a bias in our estimate due to a

non-randomly selected sample (Stock and Watson, 2003). Heckman (1979) develops a two-step statistical method (Heckit method) to correct the bias of non-randomly selected samples, which has been used widely in the estimation of wage returns to education.

Assuming our simplified wage equation is:

$$Y_j^* = \beta_1 X_{1j} + \mu_{1j} \tag{3}$$

where Y_j^* is the wage of wage earner **j**, X_{1j} is the vector of explanatory variables (education, experience, and others), and μ_{1j} is the error term of the wage equation, we can only observe Y_j^* when the person decides to be in wage employment; that is, when his actual wage is greater than his reservation wage.
The person's reservation wage, which is not observed, can be assumed to be as follows:

$$W_j^* = \gamma Z_j + \varepsilon_j$$

Where W_j^* is the person's reservation wage (not observed), and Z_j is the vector of explanatory variables (worker's personal characteristics that reflect his or her utility function and budget constraint) and ε_j is the error term of the reservation wage equation.

The person's decision to work is explained by the following rule:

$$h_j = 1 \text{ if } h_j^* \equiv Y_j^* - W_j^* > 0$$
$$h_j = 0 \text{ if } h_j^* \equiv Y_j^* - W_j^* \leq 0$$

Where h_j denotes the worker's decision either to be in wage employment $(h_j = 1)$ or not $(h_j = 0)$. We can further elaborate the selection equation as follows:

$$h_j^* \equiv Y_j^* - W_j^* = (\beta_1 X_{1j} - \gamma Z_j) + (\mu_{1j} - \varepsilon_j) = \beta_2 X_{2j} + \mu_{2j} \tag{4}$$

To sum up from the wage equation (3) and selection equation (4), we have the observation rule as follows:

$$Y_j = Y_j^*, \ h_j = 1 \qquad \text{if } h_j^* > 0$$
$$Y_j \text{ not observed, } \ h_j = 0 \text{ if } h_j^* \leq 0$$

The assumptions are:

$$\mu_1 \sim N(0,\sigma) \qquad \mu_2 \sim N(0,1)$$
$$corr(\mu_1,\mu_2) = \rho_{12} \qquad cov(\mu_1,\mu_2) = \sigma_{12}$$

From the selection equation (4), an Inverse Mills Ratio or Heckman's lambda can be calculated as:

$$\lambda_j = \frac{\phi(\beta_2 X_{2j})}{\Phi(\beta_2 X_{2j})}$$

The $\sigma_{12}\lambda_j$ is then included in the original wage equation as an explanatory variable, which contains information from those who are unemployed or do not report their earnings to explain unconditional returns to education.

$$Y_j = \beta X_j + \sigma_{12}\lambda_j + \eta_j$$

In this way, the Heckman model uses all information available to improve the estimates of the parameters in the regression model. The Heckman selection model provides consistent, asymptotically efficient estimates for all parameters in the model (Heckman, 1979; Verbeek, 2004).

Although based on a solid statistical basis, the use of the Heckit method has disadvantages. The major drawback is a formal identification problem (Verbeek, 2004). Explanatory variables used to predict the probability of being employed are also based on human capital theory. These variables, including education, labour market experience, gender, ethnicity, number of children, urban versus rural, economic situation, parents' education, and so on, also have explanatory power in the second equation. Weakly identified or exactly identified models may result in serious limitations in correcting the sample selection bias. In fact, many studies using data from different countries have shown that estimations using the Heckit method typically reveal a small sample selection bias as compared to the OLS one (Hoffman and Link, 1984; Hyclak et al., 2005; Bagheri and Kara, 2008). Certainly, one can argue that when the result of our estimation is used to determine the allocation of scarce resources or policy formulation, small differences are important, but for other purposes the sample selection bias can be minimal.

2.4. Returns to education in Vietnam

In the previous section I reviewed the literature on the estimation of returns to education using the Mincer equation and its improvements. Key limitations of the model, including endogeneity,

omitted variable bias, non-linearity of education premium, and sample selection bias were addressed.

This section will review studies on returns to education in Vietnam to ascertain how these various methods have been employed and what the findings of their application are.

Major issues in the literature on returns to education in Vietnam include rates of return, trends in rates of return during a transition period, policies inducing those changes, and income inequality between the public and private sectors. All of the studies are based on an improved Mincerian framework. The results are generally consistent, with some differences due to the different methodologies used. Another body of literature focuses on gender discrimination in the labour market, mainly based on the Blinder-Oaxaca decomposition of the gender wage gap (Liu, 2004; Liu, 2005; Pham and Reilly, 2007).

Almost all studies on the rates of return to education in Vietnam use the Household Living Standards Surveys (VHLSS 1993, 1998, 2000, 2002, 2004, 2006 and 2008) conducted by the General Statistical Office of Vietnam in collaboration with the World Bank. The survey collects information on the income and expenditure of around 9,000 households (2002-2008) and is representative at national and regional levels. VHLSS samples are designed to create unbalanced panel data for those years (especially the 2002, 2004 and 2006 dataset), which is very helpful for economic analysis.

Gallup (2002) uses two rounds of the Vietnamese Living Standards Survey (VLSS) from 1993 and 1998 and a simple OLS method to evaluate the impact of wage employment on income inequality in Vietnam. The study found that the rate of return to education in Vietnam – that is, the percentage of wage increase for one additional year in education – almost doubled from a low level of 2.9 per cent in 1993 to 5.0 per cent in 1998, but was still relatively low compared to other developing countries. Socio-economic factors included in the model, such as gender, ethnicity, region and sector, are statistically significant. Women have lower rates of return compared to men, but the gap narrowed from 31 per cent to 17 per cent during the period. There were variations in rates of return by region, with higher wage premiums in Ho Chi Minh City and Hanoi, where the education level is generally higher.

Liu (2005) uses VLSS data from 1993 and 1998 to analyse the gender wage gap in Vietnam. The empirical approach consists of OLS regressions, the Oaxaca decomposition, and separate estimates for men and women, further subdivided by sector (government, State-owned enterprises, and private sector). A vector of other socio-economic factors has been considered, including dummy variables for migrant, marital status, seven regions, urban versus rural, ethnicity, and occupation. The result shows that education is rewarded in all three sectors, with the rates of return on a one-year increase in education ranging from 3.3 per cent for men in the private sector to 7.5 per cent in the government sector. The gender wage gap is most serious in the private sector, where on average men earn 26 per cent more than women, while the gap is smallest in the government sector.

Arcand et al. (2005) focus on the use of instrumental variables (IV) in estimating the rates of return in Vietnam, using VHLSS 1993 and 1998 data. The paper analyses the validity of different types of instrument variables, from demand-side variations in schooling (parents' education, past smoking habits) to supply-side variation in schooling (proximity to school) and the Hausman Taylor (HT) matrix instrument. The result is that few instruments meet the two requirements of exogeneity and relevance. For VHLSS 1993 and 1998, only parent education and the HT matrix instrument satisfied the two conditions. IV estimates are significantly higher than OLS estimates (from 5.1 to 7.5 per cent depending on the IV, as compared to 2.6 per cent for OLS). However,

the paper is based on only 342 male wage earners observed in both 1993 and 1998, potentially resulting in a local average treatment effect (LATE) and sample selection problem.

Liu (2006) uses VLSS data from 1993 and 1998 to analyse the changes in the wages of men and women in Vietnam with different education levels during the period.

The analysis revealed that returns to an additional year of education for women increased slightly over the period while returns for men registered a decrease. This is partly explained by the fact that more women participate as wage workers in the FDI sector where education is more valued. The analysis also used Hay's two-stage method (generalization of Heckman's method) to correct sample selection bias. A multinominal logit model is used to calculate the individual probabilities of being employed in the wage sector, self-employed and unemployed. These probabilities are included as the correction term. The analysis attempts to address the identification problem by including variables correlated with labour market status but not directly with wages, such as number of children, non-labour income, and the dependency ratio. The results show that the correction term was significant for 1993 data but not significant for 1998 data. The size of the correction is arguably small.

Doan and Gibsons (2010), using VHLSS data from 1998, 2002, 2004, 2006 and 2008, analyse the trend in wage premiums over time. Their analysis shows that the rates of return to education increased significantly and consistently over the period from 3.8 per cent in 1998 to 10 per cent in 2008, contradicting the "mixed" trend reported by Liu (2006). The analysis also uses the Heckman method to correct sample section bias. Identification is achieved by including household size and household non-wage income, which affects employment probability but not wages. The estimated returns to education yielded by the model with correction are (with the exception of just one year) about one percentage point higher than those produced by the OLS model without correction. The increasing trend of returns is consistent and unchanged as compared to the OLS model.

The Phan and Coxhead (2013) study is another interesting example which not only analyses the evolution of returns to education over time, but also explains how different policies have brought about these changes. The study also relies on VHLSS data from the period 1993-2008. It is argued that both domestic policies and foreign trade liberalization have an impact on wages. While domestic policies, including SOE restructuring, first hit workers in State-owned enterprises, foreign trade liberalization first affects workers in the traded sectors. Interaction terms between wage earner categories dummies[1] with year dummies allow comparison of returns to education across groups and years. The difference-in-difference model helps to explain how domestic and foreign economic policies brought about changes in rates of return to education over the period.

The issue of an unobserved characteristic which determines workers' self-selection into different employment groups, especially the State sector, has also been raised. This results in OLS bias. The author uses the Heckman method to correct this sample selection bias. Variables relating to the dependency ratio, a household head dummy, and non-wage income are included in the employment equation, while a "network" variable (equal to 1 if there is a family member working in the State sector, and 0 if otherwise) is included in the public sector employment equation to address the identification problem. The result is that OLS estimates are biased upward, but the difference is not significant, and the trends are similar.

---

[1] These include (i) State firms and traded industries, (ii) State firms and non-traded industries; (iii) non-State firms and traded industries, (iv) non-State firms and non-traded industries

Finally, this study addresses selectivity bias due to workers' self-selection into wage employment or public employment *in a year*, but does not take into account the changing employment patterns *over a number of years*, during which workers with certain skill levels move from one sector to another, with the result that the OLS model estimates only the *conditional* rates of return to education.

Osstendorp and Doan (2013), using VHLSS from 1998 to 2006, examine the trend in returns to education during the transition period, taking into account not only the wage effect but also the employment effect of trade liberalization. It has been posited that open-door policies in Vietnam during the period not only had wage effects (changing rates of return to education), but also employment effects (changing employment structure). Previous studies on the trend in returns to education in Vietnam during the transition period using the Mincerian framework ignored the employment effect and failed to estimate the changes in the *unconditional* returns to education.

The paper starts with improving the basic Mincerian equation by including a third-order polynomial of the education variable, allowing education to have non-linear relationship with log of wage and correcting for bias induced by workers' self-selection into farming, self-employment in non-farm activities, and wage employment. Identification is achieved by including in the employment equation only variables such as household composition, marital status and household size. The result is an increasing trend in rates of return to education, especially for skilled workers. For example, from 1998 to 2006 the wage premium for high school graduates increased from 3 per cent to 5.4 per cent, while the wage premium for 15 years of education increased from 7.4 per cent to 13.3 per cent. This is in line with earlier studies on the trend in returns to education in Vietnam using the same data for the period.

The paper goes further in pointing out that the labour market is segmented and that returns to education and sample selectivity term differ across sectors, suggesting that the change in employment structure would create a bias in the estimation of rates of return to education over time. The unconditional return to education is decomposed into wage effect and employment effect. The conclusion is that the changing employment patterns consequent on trade liberalization and economic reform result in the Mincerian returns significantly overestimating the unconditional returns to education.

In summary, as data has been accumulated over the years, more and more studies have been conducted examining the causal linkage between education and labour market outcomes, and its evolution over the transition period. Most of the results are consistent, revealing a low but increasing trend in rates of return. However, the exact rates of return slightly differ between studies, owing to differing techniques. An updated analysis is required in order to shed light on this issue.

3. Data and Methodology

The previous chapter reviewed literature on the estimation of returns to education based on human capital theory. It has been noted that the basic Mincerian model suffers from a number of key limitations, including endogeneity of individual ability, omitted variable bias, and sample selection bias. Despite these weaknesses, a number of techniques can be applied to address the problems. The chapter also reviewed key studies on returns to education in Vietnam.

This chapter will discuss the data available and identify which methodology will be applied to analyse the returns to education in Vietnam.

3.1. The Labour Force Survey 2012

In the past, most studies on returns to education in Vietnam have been based on the Vietnam Household Living Standard Surveys. This paper is based on panel data for a high-frequency Labour Force Survey (LFS) in 2012. The survey was conducted in 2011 by the General Statistical Office (GSO) on a quarterly basis with technical support from the International Labour Organization. The main objective of the survey was to collect data on key indicators concerning Vietnam's labour market, including those employed, those unemployed, and those not in the labour force, and to provide descriptive and explanatory data on each category.

The LFS 2012 uses stratified random sampling methodology. Each province or city under the direct management of the government constitutes the first layer, with two sub-layers: rural and urban. Enumeration areas are chosen using the Kish selection methodology[2] for the list of enumeration areas, drawn from 15 per cent of the 2009 Population and Housing Census. Within each enumeration area, 20 households were selected at random. All persons aged 15 years or over were eligible to be interviewed. In total 202,520 households were visited and 746,768 people were interviewed in 2012.

The Labour Force Survey in Vietnam used a rotating sample design, 2-(2)-2 pattern. A given household was included in the sample for two consecutive quarters, was omitted for two quarters and re-included for another two consecutive quarters before being dropped out of the sample forever. Probability weights for each observation were provided. Data for each quarter was representative at national and regional level. Households interviewed in the first quarter and in the third quarter were re-interviewed in the second and fourth quarters respectively. Data for the year is representative at provincial level.

Key information collected in the LFS 2012 included:

   − Household and personal characteristics: family size, relationship with household head, region, urban versus rural, gender, date of birth, ethnicity, migration, schooling status, schooling attainment, marital status;

   − Employment-related information: employment status, additional part-time jobs, wage, job-seeking, informality, earning from main job, earning from part-time jobs, hours of work, underemployment, economic sector, ownership type, social security, firm size, employment contract.

3.2. Methodology

The panel data provided by the 2012 Labour Force Survey offers a very good opportunity to revisit the causal linkage between education attainment and labour market outcomes for wage earners in Vietnam. The survey questionnaire was developed with the technical support of the International Labour Organization, and contained most of the explanatory variables needed to explain variations in earnings on the basis of human capital theory and the Mincerian framework.

---

[2] Leslie Kish described a rigorous, almost pure probability method of sampling persons within households to be surveyed. He named two basic conditions: 1) there must be a known probability of inclusion (excluding zero) of each adult in the population, and 2) it must be a practical and efficient procedure to implement (SAGE, Research Methods).

This section discusses the advantages and disadvantages of using the Labour Force Survey to analyse returns to education.

Having obtained microdata and the manual for the 2012 LFS, standard ILO definitions have been used to construct key employment variables, including those employed, unemployed or inactive[3] and informal employment[4] categories. As recommended, this paper adopts a relaxed definition of unemployment.[5] National statistical classification tables have been used to construct key variables, including regions[6] (by level of economic development) and economic sector.[7]

The 2012 LFS reports only the level of education attainment, which is primary, lower secondary, high school, college or above. It does not report the exact number of years of schooling that the person has completed. This may result in our model slightly overestimating the rates of returns to schooling. However, as argued earlier, the "sheepskin effect", under which employers value completed qualifications over non-credentialed years, suggests that the problem is not significant. As in other studies, education has been divided into 5 categories, (i) *no education*: the individual has never been to school or has not completed primary education; (ii) *primary*: the individual has completed five years of education; (iii) *lower secondary*: the individual has completed nine years of education; (iv) *upper secondary*: the individual has completed 12 years of education; and (v) *college and above*: the individual has completed 15 years of education or more. A dummy variable has been created for each category to estimate how big an increase in wage would result from a change from no education into each education category. Age is used as a proxy for labour market experience.

The extended panel Mincerian regression will be:

$$Y_{it} = \beta_0 + \beta_1 X_{it} + \beta_2 Z_{it} + \beta_3 Z_{it}^2 + \beta_4 T_{it} + \mu \tag{5}$$

Where: Y = log of hourly wage

X = education level dummies

Z = age

T = a vector of other explanatory and controlling variables (gender, economic sector, type of ownership, migrant, region, urban, etc)

The 2012 LFS offers very limited scope for using instrumental variables to address endogeneity problems due to the unobserved ability of wage earners. It has been argued that, in Vietnam, parent education is one of the few instruments that have been found to meet the conditions for

---

[3]   ILO   Statistical   Resolution: http://www.ilo.org/wcmsp5/groups/public/---dgreports/---stat/documents/normativeinstrument/wcms_087481.pdf

[4]   ILO   Guidelines   concerning   a   statistical   definition   of   informal   employment: http://www.ilo.org/public/english/bureau/stat/download/guidelines/defempl.pdf

[5] Three are the conditions required to define an unemployed: 1) without work, 2) currently available for work, 3) seeking for work. The 13th ICLS introduced a provision which allows for relaxation of the seeking work condition (Hussmanns, 2007).

[6] General Statistical Office: provincial code and classification of provinces into 6 regions in Vietnam: http://www.gso.gov.vn/khodulieuldvl/MetaData.aspx?Mct=3&NameBar=SI%C3%8AU%20D%E1%BB%AE%20LI%E1%BB%86U%20%3E%20Kh%C3%A1i%20ni%E1%BB%87m,%20%C4%91%E1%BB%8Bnh%20ngh%C4%A9a,%20c%C3%A1ch%20t%C3%ADnh

[7]   General   Statistical   Office: classification   and   codes   of   economic   sectors   in   Vietnam http://www.gso.gov.vn/Modules/Doc_Download.aspx?DocID=11339

instrumental variable as explained in Section 2.3. Unfortunately, the data on education is not fully captured by the survey. There is no question on father's or mother's education, although there is a question on the relationship with the household head, from which information about father's or mother's education can be obtained if the mother or father is living in the same household as the wage earner. However, in general, young workers in richer families can afford the high accommodation expenses incurred by living separately from their parents, while those in poorer families have to live with their parents. Therefore, using the instrument with information on the education of the parents currently living in the same household as the wage earner would introduce an additional selection bias to the model.

The data reveals that there is a potential sample selection bias due to workers' self-selection into wage employment. A typical characteristic of a developing country's labour market is a large informal sector where most workers have no stable income or do not report their income. The 2012 LFS shows that only 34.8 per cent of the workforce (17.9 million out of 51.4 million workers) are wage earners. Almost all (91 per cent) are in the formal sector. This underscores the importance of examining the magnitude of the sample selection bias.

The regression was run with the Heckman correction model to examine the sample selection bias. A number of variables – including dummies for being married, being the household head, household size, and number of natural children – are included in the selection equation to explain workers' decisions to be in wage employment. The rationale is that employers are unlikely to discriminate against a person based on grounds of how many persons there are in his or her family or whether he or she is the head of a household, whereas this information does influence a worker's decision to work. Table A1 in Appendix A, model (4), shows the detailed regression results. All the regressions are significant (Prob>chi2 =0.000). Figure 1 below summarizes the difference in rates of return obtained with and without the Heckman selection model. The results show that OLS consistently overestimates rates of return to education because of the selection bias. However, the bias is not large for any of the estimates, standing at a maximum of around 10 per cent. This is in line with earlier studies on returns to education in Vietnam (Doan and Gibsons, 2010; Phan and Coxhead, 2013).
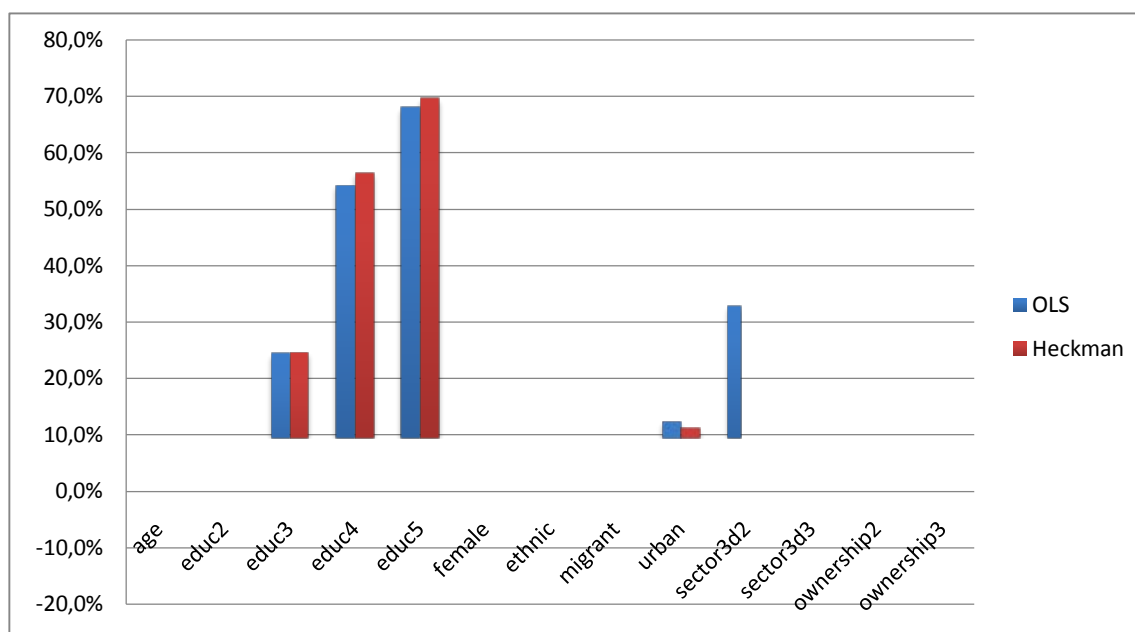
Figure 1: Differences between OLS and Heckman corrected estimates

Source: Vietnam Labour Force Survey 2012.

To further test the results of the analysis, a panel fixed effects model was also run. Because of the limited variability of the variables over time, it was not possible to recover significant results for most of them. Nonetheless, the basic results for the returns to education are broadly confirmed.

In summary, while recognizing the potential problems of endogeneity and sample selection bias, the scope for addressing these issues with the available data from the 2012 LFS is limited. Interpretation of the coefficients or comparisons with other studies should therefore be made with caution. However, the panel nature of the LFS 2012 still offers a very good opportunity for examining the rates of return in Vietnam's labour market, and a good reference for comparison with future studies using the same data.

4. Labour Market Returns to Education

The previous chapter discussed fundamental issues in methodology and proposed a model for examining rates of return to education in Vietnam. This chapter sets out the results of the analysis. First, descriptive statistics are presented, followed by a more detailed discussion of returns to education in terms of labour market status and earnings by economic sector, ownership, and gender.

4.1. Descriptive statistics

As in other developing countries, Vietnam's labour market is characterized by a low unemployment rate, but a high percentage of working poor with low incomes, notwithstanding long working hours. In 2012, of the 68.2 million people aged 15 and over, 51.4 million were employed and 938,000 unemployed. The workforce participation rate was 76.8 per cent and the unemployment rate 1.8 per cent. The low unemployment rate of 1.8 per cent is explained by the fact that poor people cannot afford to stay idle and have to work, usually in the informal sector with a meagre income.
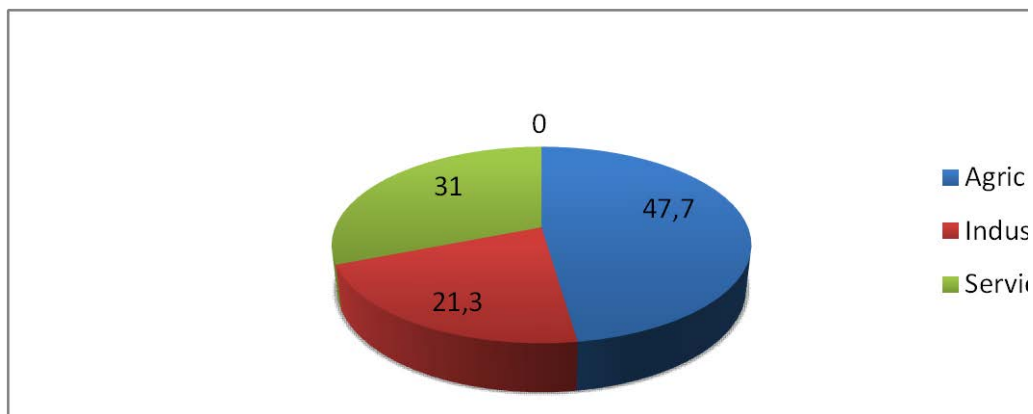
Figure 2: Share of total employment by sector (in %)

Source: Vietnam Labour Force Survey, 2012.

Agriculture is an important sector, accounting for 47.7 per cent of total employment, but contributing only 20 per cent of GDP. Meanwhile, industry, which accounts for 21.3 per cent of total employment, contributes 41 per cent of total economic output. In 2012, 70 per cent of employed workers were in rural areas.

Education

The educational attainment level of the workforce is generally low. The majority (69 per cent) of the working-age population have only primary education, while those with tertiary education account for 5.7 per cent. Higher-educated workers (upper secondary and tertiary level) are more likely to be unemployed (4.3 per cent and 2.4 per cent compared to 0.7 per cent, 1 per cent and 2.1 per cent of the lower levels), but they are less likely to stay inactive (around 12 per cent as compared to 22.5 per cent, 26.2 per cent and 37.6 per cent of the lower levels). The higher unemployment rate can be explained by the fact that educated people have higher reservation wages. In addition, it is important to underline two other possible explanations. First, the structural mismatch between education and availability of jobs: the demand side of the labour market is not able to absorb the skills and knowledge of part of graduates. The second factor is the low quality of education.
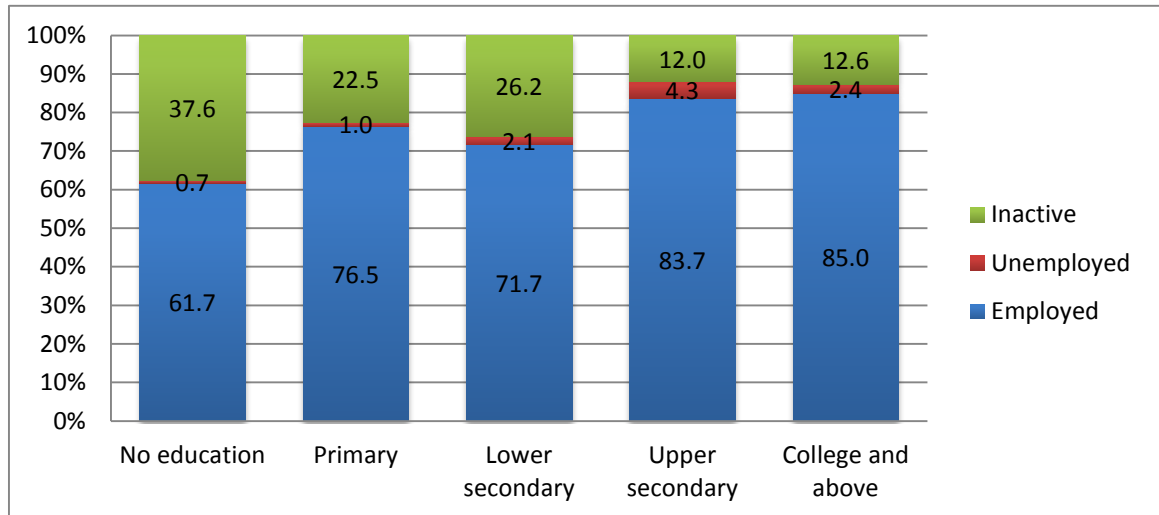
Figure 3: Employment status by educational levels (in %)

Source: Vietnam Labour Force Survey, 2012.

More women than men have received tertiary education (56.14 per cent as compared to 43.86 per cent). Indeed, the differences between men and women at each education level are quite big. This suggests that it may be a good idea to insert interaction dummies between gender and education levels so as to be able to estimate the returns to education separately for each group.

The economic sectors that absorb most of highly educated workers include information technology, financial services, the public sector, professional services, and education, whereas workers with low education tend to work in households, accommodation and catering, transportation, vendoring, construction, manufacturing and, mostly, in agriculture. In general, the service sector absorbs more educated workers than do either industry or agriculture. The public sector attracts the most educated part of the workforce, while most workers with lower levels of education take low-end jobs in the private sector.
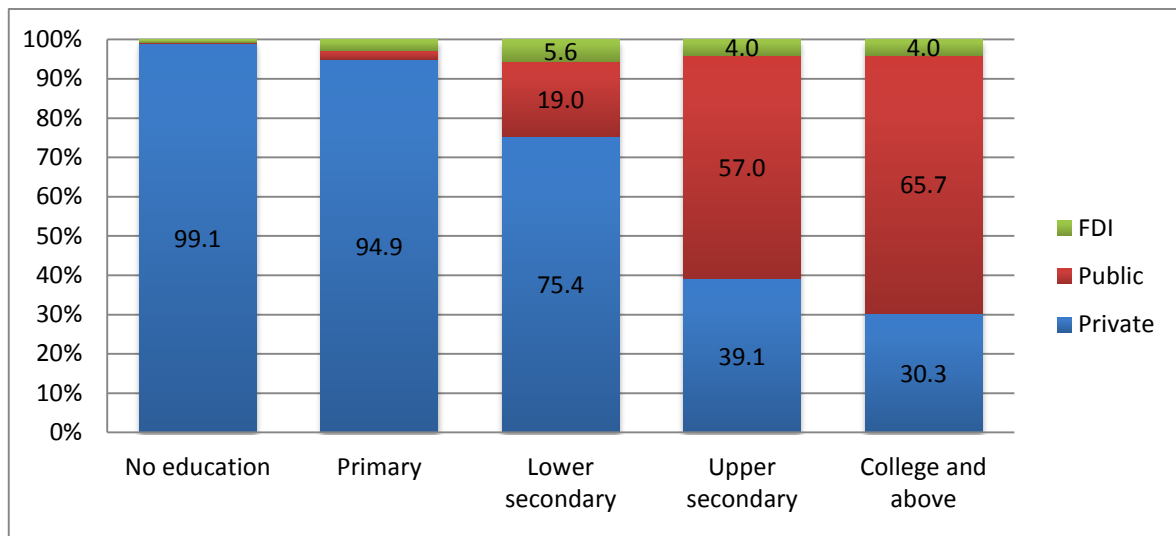
Figure 4: Distribution of education by types of ownership (in %)

Source: Vietnam Labour Force Survey, 2012.

16

Earnings

On average, workers with higher education levels are paid much more per hour than those with lower levels of education. Male university graduates earn 37,300 VND/hour (US$1.77), 2.8 times more than those who did not attend school or did not complete primary schooling. Earnings by education level appear to be convex, where the wage premium at the higher levels of education is higher than at lower levels. Men are consistently paid more than women with the same education level.
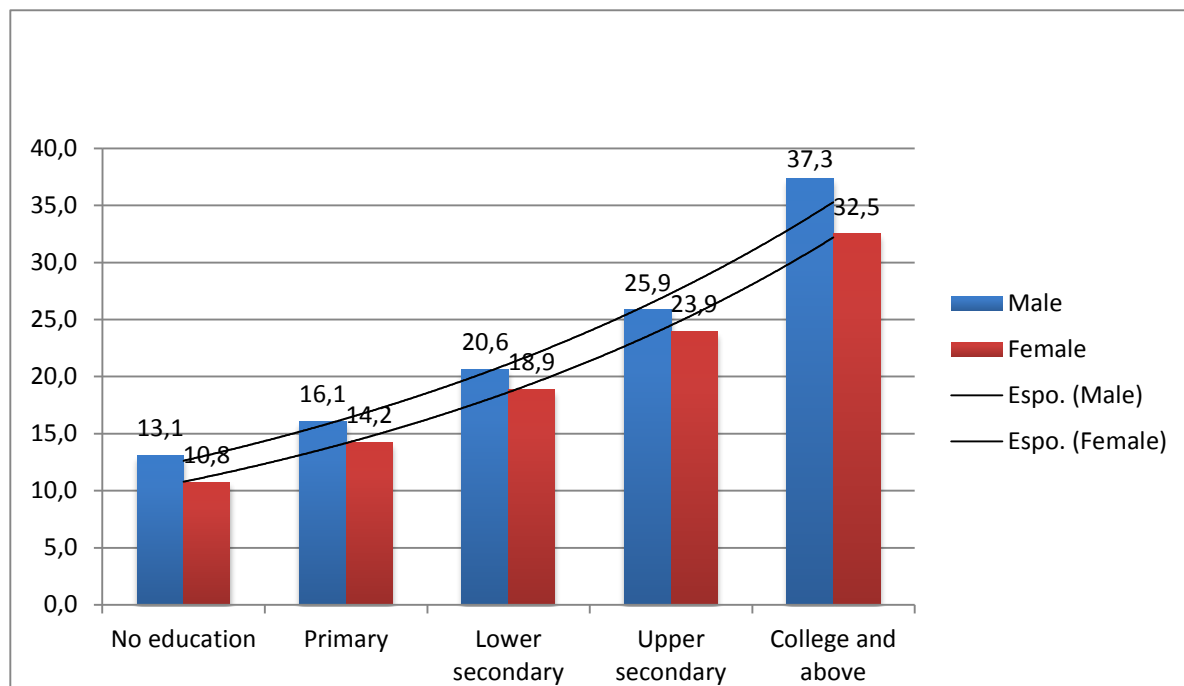
Figure 5: Average hourly income by education and sex (in thousand VND)

Source: Vietnam Labour Force Survey, 2012

The best paid sectors (hourly wage) are finance and insurance, and education. Agriculture, catering and accommodation are the worst paid. In general, although the service sector is more heterogeneous, ranging from very low-end services in the informal sector such as vendoring and motorcycle taxis to the best paid sectors such as banking, finance, insurance, and education, workers in this sector are still, on average, better paid than those in industry, while workers in industrial sectors are better paid than those in agriculture.

Figure 6 shows that the wage premium for education attainment across types of ownership varies, with a much higher wage premium offered by the FDI sector.

Although the public sector employs the bulk (65.7 per cent) of the most educated workforce in Vietnam, it can only pay college or university graduates an hourly wage at a level of 63 per cent of that offered by the FDI sector.
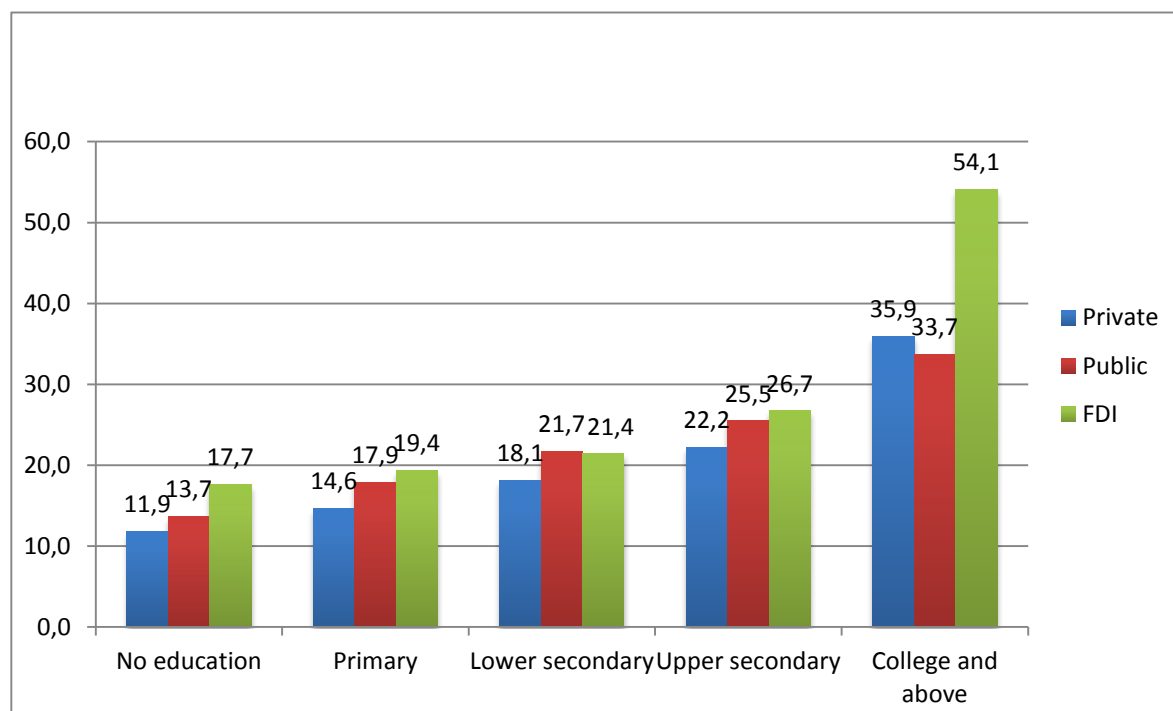
Figure 6: Average hourly income by education and ownership (in thousand VND)

Source: Vietnam Labour Force Survey, 2012

Data reveals that other characteristics of workers and enterprises may also help to explain variations in earnings. For example, workers with more than two children earn less than workers with two, one or no children. On average, migrant workers are better educated and earn more than local workers. And workers in an ethnic minority group are consistently lower-paid at all education levels.

4.2. Returns to education in Vietnam

Although descriptive statistics show interesting findings, one needs to make allowances for covariates when examining the statistical relationship between education and earnings.

Random effect estimations over the panel data were run. The results are reported on in Appendix A, Table A1, A2, and A3.

Table A1 compares the results of the basic Mincer model, improved Mincer models and the Heckman corrected model. As usual in this type of analysis, the first version of the model includes only the basic Mincer predictors, such as age and level of education. These simple estimations are not reliable, so in a second and third step, personal characteristics and other control factors are added. All coefficients in the basic and improved Mincer models are significant at 1 per cent, except for variables on children (at 5 per cent and 1 per cent) and married individuals (not significant at 10 per cent). It is evident that the basic Mincer model overestimates the returns to education. Higher R-squared values in the model with personal characteristics (2) and the model with other controlling variables (3) imply that the variables included have significant explanatory power over the variations in earnings.

Taking model (3) as reference, one can extract some interesting findings. First, an increase of one year of age is associated with a wage increase of approximately 5.04 per cent (i.e. 5.21 per cent minus 2*0.07*age). The coefficient of age squared is negative, as expected, showing a concave relationship between age and earnings. This means that, on average, workers' wages start to decrease at the age of 37.2. Table A3 shows that the corresponding turning point for industry and service workers is higher, at the age of 41.2 and 42.6 respectively.

The analysis of the return to education by educational level shows that it is 4 per cent for the primary level, 24.6 per cent for lower secondary, 54.4 per cent for upper secondary and 68 per cent for college or university. More specifically, on average, one additional year of primary education increases wages by 0.8 per cent, one additional year of lower secondary by 2.7 per cent, one additional year of upper secondary by 4.5 per cent, and one additional year of university by 4.5 per cent. In general, having a university degree does not add an extra premium per year as compared to the upper secondary level. In other words, continuing a year of higher education gives a worker a lower wage increase compared to the rates of return for each year of secondary education, while the costs for tertiary education are much higher. This may be due to low demand for skilled workers in some sectors such as agriculture or, more likely, to the mismatch between supply and demand of work or the low quality of tertiary education.

As expected, having a high education level is more significant in the formal sector. In this sector, a university graduate will be paid 80 per cent more. The return to lower levels of education is not significant and sometimes even negative in the informal sector (Table A2). This implies that salaries in the formal sector are higher and in line with workers' productivity, while this is not true of the informal sector.
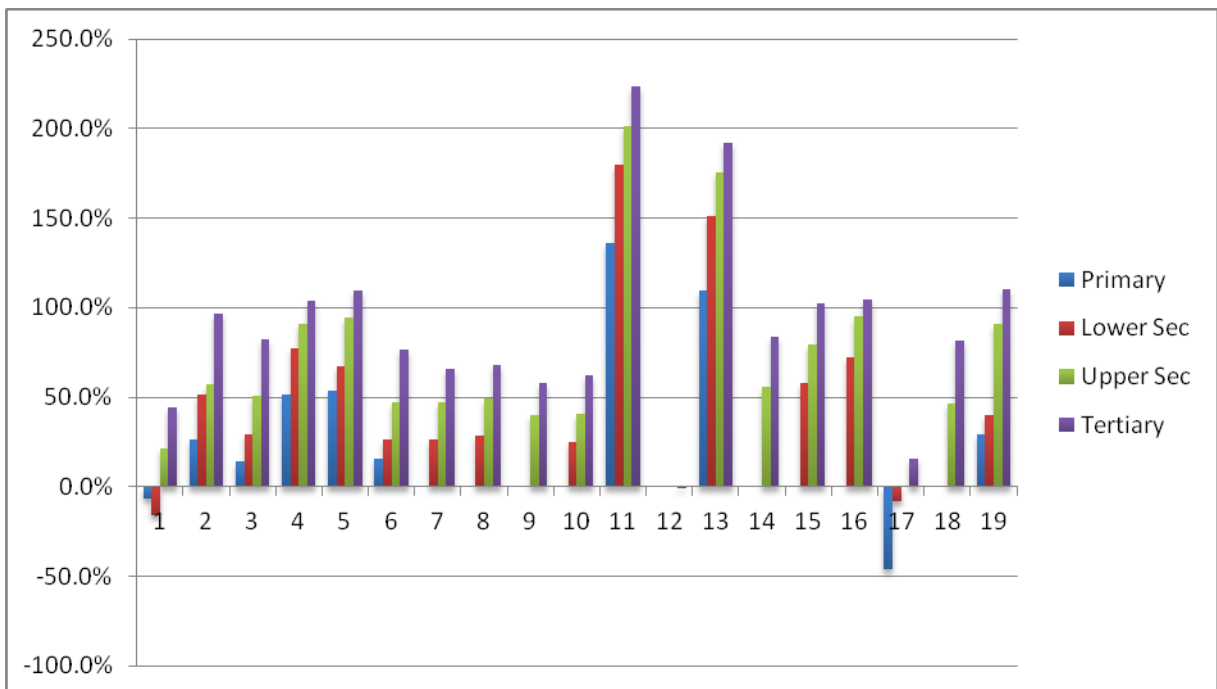

Figure 7: Returns to education by economic sector (in %)

Source: Vietnam Labour Force Survey, 2012.
Note: 0 includes not statistically significant cases
All cases are significant at 10% or higher

19

Figure 7 reveals the two most notable subsectors in which education is highly rewarded. These are banking, finance and insurance (columns 11) and science, technology and professional activities (columns 13), which respectively pay 224 per cent and 190 per cent more to college and university graduates as compared to a worker with no education. Meanwhile, in real estate services, health and social services, and agriculture, investments in education – especially at the lower levels – do not seem to pay off.

All else being equal, women earn on average 15.6 per cent less than their male counterparts (Table A1). This gender wage gap is particularly high in some economic subsectors such as real estate services, 37 per cent; mining, 28.3 per cent; transportation and storage, 27.4 per cent; and wholesale and retail trade, 27.1 per cent (Figure 8). These subsectors make the gender wage gap most significant in industry (26 per cent), while that in agriculture and services is smaller (16 per cent and 9 per cent respectively). The gender pay differential also varies by type of ownership. It is only 2.7 per cent in the public sector, but as high as 13.8 per cent for FDI and 26.7 per cent in the domestic private sector (Table A2).

In general, women have higher rates of return across all education levels. The higher the education level, the smaller the gap (Table A2). This concurs with the findings of earlier studies on returns to education in Vietnam (Liu, 2006).
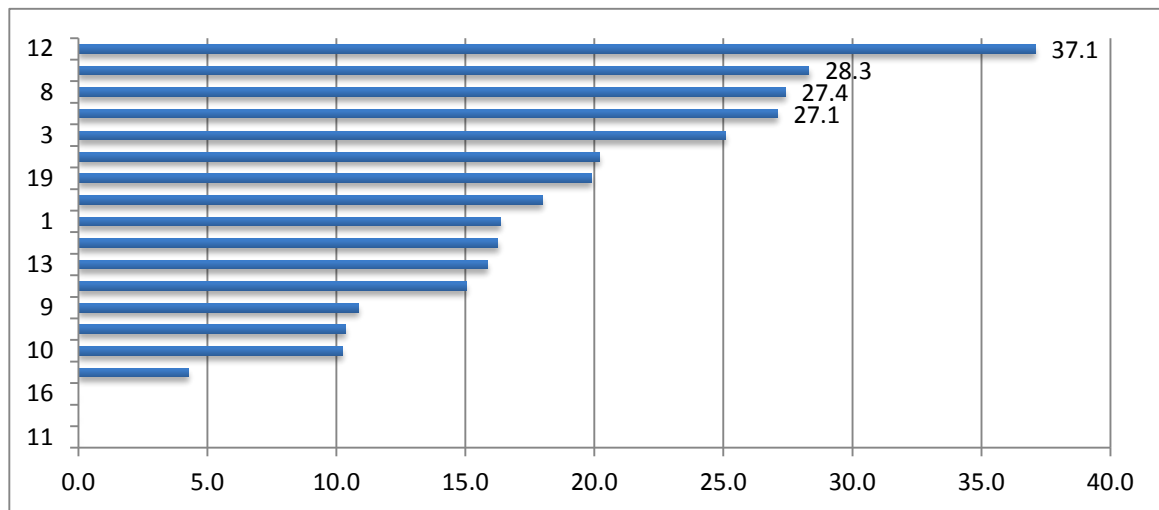
Figure 8: Gender wage gap by sector (in %)
Source: Vietnam Labour Force Survey, 2012.
Note: 0 includes not statistically significant cases
Gender wage gap of sector 17 is significant at 10%, all other non-zero cases are significant at 1%

Another source of gender discrimination is the presence of children. Looking at the reference regression outcome (regression 3 in Table A1), it is possible to see that, for having an additional child, a worker earns from 0.8 per cent to 1.5 per cent less, depending on whether the child is an infant, does not go to school, is attending primary school or is a grown-up (Table A1). However, disaggregating the regression by gender shows that having children mostly affects women in the formal sector, where the wage is consistently reduced by between 2.5 and 4.3 per cent for every additional child (Table A2). The result for the informal sector is insignificant or sometimes mixed. This is perhaps due to the fact that in Vietnam it is commonly perceived that women are in charge of taking care of children. Employers would therefore expect that women with children would not contribute to work as much as when they are single or have grown-up children. Men

are free from this assumption, suggesting that this wage reduction can be attributed more to gender discrimination than to a genuine economic reason.

In contrast to earlier studies finding that marital status has a sizable effect on wages (Pham and Reilly, 2007; Liu, 2004; Brassard, 2004), and suggesting the existence of gender discrimination for this indicator, this study shows that marital status does affect labour market status, but it does not impact significantly on wages (Table A1). This holds true when disaggregating the regression by gender (Table A2) and by sector (Table A3), except for agriculture and the informal sector where a worker will earn significantly less (around 10 per cent) if they are married.
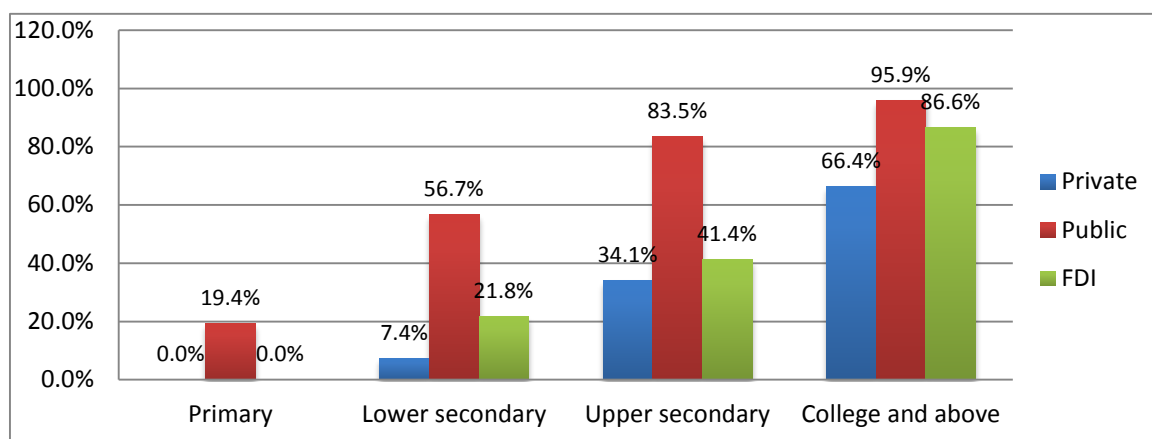
Figure 9: Returns to education by types of ownership

Source: Vietnam Labour Force Survey, 2012.

Note: 0 includes not statistically significant cases

All other cases are significant at 1%

Figure 9 shows the returns to education by type of ownership. There is no evidence that the private sector and FDI sector value primary education, in contrast with a strong wage premium in the public sector for workers with primary education. This is explained by the fact that in the public sector workers' salaries are paid according to a fixed wage scale and wage table, where each level of education is rewarded with a fixed wage premium across the sector, whereas salaries in the private and FDI sectors are based on negotiation between workers and employers. Higher returns to education in the public sector imply that the wage scale and wage table of the public sector overestimate the market value of education, especially the lower levels (primary, lower secondary and upper secondary levels).

Higher returns to education in the FDI sector, as compared to the private sector, show a stronger demand for educated workers. The actual hourly wage for high skilled workers in the FDI sector is much higher than in both the public and private sectors (Figure 6). This implies that the FDI sector has a comparative advantage in attracting skilled human resources because of both its willingness and its ability to pay.

Finally, it is interesting to analyse the condition of migrant workers. They earn, on average, around 5 per cent more than non-migrants. They are also better educated than resident workers. This implies that domestic migrants in Vietnam have access to both better education and economic opportunities. It is important to note that, in the formal sector, male migrants are much better off than female migrants: the former earn up to 6.33 per cent more than their

resident counterparts while the latter only earn 2.4 per cent more than their resident counterparts (Table A2).

5. Conclusion

The aim of this paper was to provide an updated analysis of the returns to education in Vietnam using the 2012 National Labour Force Survey. The panel nature of the national survey provides very reliable results which are, in general, in line with earlier studies on returns to education using the Vietnam Household Living Standards Surveys which applied the same methodology.

It has been argued that in Vietnam the basic Mincerian equation overestimates the returns to education owing to omitted variable bias. When including all the controlling factors, the model still overestimates the actual rates of return owing to sample selection bias. However, the Heckman correction model shows that the size of the bias is not large (no more than 10 per cent) in most cases.

The study shows that the returns to education are still low, especially for tertiary level, indicating either low demand for skilled workers or, more likely, the mismatch between education and job availability, and the low quality of tertiary education. Investments in higher levels of education in a number of sectors such as real estate services, health and social services and agriculture do not pay off.

A gender wage gap is present in almost all economic sub-sectors. There are also indications of wage discrimination based on the gender-related role of caring for children. However, when the number of children is taken into account, marital status does not have a significant impact on wages.

The study also shows different levels of returns to education by types of ownership, where the wage scales and wage tables provided by the government for the public sector overvalue the market rates of return to education. The FDI sector has a comparative advantage in recruiting educated workers because of both its willingness and its ability to pay.

In summary, the panel nature of the 2012 National Labour Force Survey offers a very good opportunity for revisiting the causal linkage between education attainment and labour market outcomes for wage earners in Vietnam. This paper offers a very good starting point for analysing the trends in rates of return to education in this country using the quarterly survey. As the employment structure in Vietnam continues to change – and as suggested by previous papers – future studies on the trends in returns to education in Vietnam should take into account both the wage effect and the employment effect of trade liberalization.

References

Angrist, J. D. and Krueger, A. B. (1991). 'Does Compulsory Schooling Attendance Affect Schooling Decisions?', *Quarterly Journal of Economics*, 106 (4),pp. 970-1014.

Arcand, J., D'Hombres, B. and Gyselinck, P. (2005). 'Instrument Choice and the Returns to Education: New Evidence from Vietnam', Labor and Demography 0510011, EconWPA, Last accessed on 26 May 2014 at http://ideas.repec.org/s/wpa/wuwpla.html

Bagheri, F. and Kara, O. (2008). 'Selection Bias and the Estimated Rate of Return to Education', *Review of. Business Research*, 8, pp. 24-35.

Belzil, C. and Hansen, J. (2002). 'Unobserved Ability and the Return to Schooling', *Econometrica*, 70, pp. 575-591.

Belzil, C. (2006). 'Testing the Specification of the Mincer Wage Equation', Working Paper 0608, Groupe d'Analyse et de Théorie Economique (GATE), Centre national de la recherche scientifique (CNRS), Université Lyon 2, Ecole Normale Supérieure.

Becker, G. S. (2008). '*Human Capital',* last accessed on 26 May 2014 at http://www.econlib.org/library/EncHumanCapital.html.

Brassard, C. (2004). 'Wage and Labour Regulation in Vietnam within the Poverty Reduction Agenda', Public Policy Programme, National University of Singapore.

Card, D. (1999). 'The Causal Effect of Education on Earnings', in Ashenfelter O. and Card D. (Eds.). *Handbook of Labor Economics*, New York, North-Holland.

Card, D. (2001). 'Estimating the Returns to Schooling: Progress on Some Persistent Econometric Problems', Econometrica, 69, pp. 1127-1160.

Deschenes, O. (2001). 'Unobserved Ability, Comparative Advantage and the Rising Return to Education in the United States: A Cohort-Based Approach', Princeton University Industrial Relations Section Working Paper No 465.

Doan, T. and Gibson, J. (2010). 'Return to schooling in Vietnam during economic transition: does the return reach its peak?', MPRA Paper No24984, Economics Department, University of Waikato.

Gallup, J. L. (2002). 'The wage labor market and inequality in Viet Nam in the 1990s', World Bank Policy Research Working Paper No. 2896.

Griliches, Z. (1977). 'Estimating the Returns to Schooling: Some Econometric Problems', *Econometrica*, 45 (1), pp. 1-22.

Harmon, C., Oosterbeek, H. and Walker, I. (2003). 'The Returns to Education: Microeconomics', *Journal of Economic Surveys*, 17 (2), pp. 115–156.

Heckman, J. J. (1979). 'Sample selection bias as a specification error', *Econometrica,* 47 (1), pp. 153-161.

Heckman, J. J. and Urzua, S. (2009). 'Comparing IV with Structural Models: What Simple IV Can and Cannot Identify', National Bureau of Economic Research Working Paper 14706.

Heckman, J. J., Lochner, L. J. and Todd, P. E. (2003). 'Fifty Years of Mincer Earnings Regressions', last accessed on 26 May 2014 at http://time.dufe.edu.cn/mingrendt/lochner030404.pdf.

Hoffman, S. D. and Link, C. R. (1984). 'Selectivity Bias in Male Wage Equations: Black-White Comparison', *The Review of Economics and Statistics*, 66 (2), pp. 320-324.

Hussmanns, R. (2007). 'Measurement of employment, unemployment and underemployment-Current international standards and issues in their application', ILO Bureau of Statistics, Geneva.

Hyclak, T., Johnes, G. and Thornton, R. (2005). *Fundamentals of Labor Economics*, Boston, Houghton Mifflin Company.

Lall, A. and Sakellarious, C. (2010). 'Evolution of Education Premiums in Cambodia: 1997-2007', *Journal of the East Asian Economic Association*, 24 (4), pp. 333-354.

Lam, D. and Schoeni, R. (1993). 'Family background, education and employment in urban Ethiopia', *Oxford Bulletin of Economics and Statistics*, 1, pp. 167-183.

Liu, A. Y. C. (2004). 'Gender Pay Gap in Vietnam: 1993-1998', *Journal of Comparative Economics,* 32 (3), pp. 586—596.

Liu, A. Y. C. (2005). 'Impact of economic reform on wage structure in Viet Nam', *East Asian Economic Perspectives,* 16 (2), pp. 46–57.

Liu, A. Y.C. (2006). 'Changing wage structure and education in Vietnam 1993–1998: the roles of demand', *Economics of Transition,* 14 (4), pp. 681–706.

Lemieux, T. (2003). 'The Mincer Equation Thirty years after Schooling, Experience and Earnings', Working Paper No. 62, Center for Labor Economics, UC Berkeley, last accessed on 26 May 2014 at: http://emlab.berkeley.edu/users/cle/wp/wp62.pdf.

Flabbi, L., Paternostro, S. and Tiongson E. R. (2008). 'Returns to education in the economic transition: A systematic assessment using comparable data', *Economics of Education Review*, 27 (6), pp. 724-740.

Mincer, J. A. (1974). Schooling, Experience, and Earnings. NBER (distributed by Columbia University Press), New York.

Mincer, J. A. (1997). 'Changes in Wage Inequality, 1970-1990', *Research in Labor Economics*, 16, pp. 1-18.

Ministry of Education and Training (MOET) (2014). Current Situation of Vietnamese Education, last accessed on 26 May 2014 at: http://en.moet.gov.vn/?page=6.1&view=3451.

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

Phan, D. and Coxhead, I. (2013). 'Long-run costs of piecemeal reform: Wage inequality and returns to education in Vietnam', *Journal of Comparative Economics,* 41 (4), pp. 1106-1122.

Pham, H. T. and Reilly, B. (2007). 'The gender pay gap in Vietnam, 1993–2002: a quintile regression approach', *Journal of Asian Economics,* 185 (5), pp. 775–806.

Stock, J. H. and Watson, M. W. (2010). *Introduction to econometrics: 3rd ed.* Boston, MA: Pearson Education Inc.

Verbeek, M. (2004). *A Guide to Modern Econometrics: 2nd edition*, John Wiley & Sons.

World Bank (2008). Vietnam: Higher Education and Skills for Growth, Washington, DC. last accessed on 26 May 2014 at https://openknowledge.worldbank.org/handle/10986/7814.

World Bank (2014). Vietnam Overview, last accessed on 26 May 2014 at http://www.worldbank.org/vi/country/vietnam/overview.

Appendix A: Results of the estimations

Table A1: Returns to education in Vietnam by random effects estimation

VARIABLES	(1) Basic Mincer model	(2) With personal characteristics	(3) With all control factors	(4) With Heckman correction Wage equation	Selection equation
Age	0.0456***	0.0479***	0.0521***	0.0559***	
	(48.67)	(26.13)	(29.72)	(38.54)	
age2	-0.0006***	-0.0006***	-0.0007***	-0.0007***	
	(-46.95)	(-29.51)	(-32.71)	(-41.60)	
Education dummies: No education is the reference					
Primary level	0.2747***	0.1612***	0.0396***	0.0327**	0.4184***
	(23.94)	(11.11)	(2.79)	(2.49)	(22.46)
Lower secondary	0.5196***	0.4823***	0.2459***	0.2474***	0.7247***
	(44.19)	(32.06)	(16.19)	(17.13)	(36.85)
Upper Secondary	0.7639***	0.8207***	0.5440***	0.5655***	1.2680***
	(58.46)	(48.16)	(31.14)	(33.07)	(46.69)
College and above	1.0020***	0.9908***	0.6812***	0.6973***	1.3116***
	(83.71)	(64.83)	(42.66)	(44.24)	(61.87)
female		-0.1477***	-0.1559***	-0.1510***	
		(-28.37)	(-31.28)	(-36.59)	
child06		-0.0287***	-0.0087**	-0.0170***	
		(-6.76)	(-2.14)	(-4.30)	
child612		-0.0290***	-0.0087**	-0.0226***	
		(-7.29)	(-2.27)	(-5.48)	
child1215		-0.0389***	-0.0137***	-0.0268***	
		(-7.28)	(-2.66)	(-4.93)	
ethnicity		-0.1921***	-0.0966***	-0.0882***	
		(-25.85)	(-12.33)	(-13.43)	
married		-0.0087	-0.0048	-0.0198**	0.7058***
		(-0.98)	(-0.56)	(-2.54)	(62.33)
migrant		0.0725***	0.0344***	0.0494***	
		(17.29)	(8.26)	(12.24)	
urban			0.1241***	0.1123***	
			(23.24)	(26.07)	
Sector dummies: agriculture is the reference					
Industry			0.3299***	0.3254***	
			(47.09)	(50.41)	
Services			0.2465***	0.2136***	
			(32.12)	(30.38)	
Region dummies: Hanoi is the reference					
region1			-0.1629***	-0.1481***	
			(-15.00)	(-17.34)	
region2			-0.1689***	-0.1628***	

			(-15.04)	(-18.53)	
region3			-0.1714***	-0.1604***	
			(-16.06)	(-19.17)	
region4			-0.1228***	-0.1133***	
			(-9.89)	(-11.49)	
region5			0.1047***	0.1063***	
			(9.04)	(11.71)	
region6			-0.1824***	-0.1637***	
			(-16.17)	(-18.42)	
region8			0.1578***	0.1649***	
			(12.46)	(16.68)	
Types of ownership dummies: domestic private sector is the reference					
Public			0.2275***	0.2318***	
			(36.29)	(41.53)	
FDI			0.1505***	0.1728***	
			(14.25)	(18.69)	
Natural children					0.6696***
					(126.62)
Household size					-0.0865***
					(-25.94)
Household head					0.3883***
					(47.93)
Constant	1.4292***	1.5970***	1.3429***	1.3319***	-1.3470***
	(70.21)	(38.92)	(32.86)	(37.19)	(-60.71)
Observations	147,991	88,142	86,915	149,323	149,323
Number of id	91,721	52,924	52,340		
R-squared	0.254	0.308	0.373		

z-statistics in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Source: Vietnam Labour Force Survey, 2012

Table A2: Returns to education by random effects estimation – Male vs. female

VARIABLES	(1) Male	(2) Female	(3) Formal &male	(4) Formal &female	(5) Informal &male	(6) Informal &female
Age	0.0538***	0.0414***	0.0552***	0.0428***	0.0244***	0.0327***
	(21.89)	(16.11)	(23.72)	(17.56)	(3.87)	(3.84)
age2	-0.0007***	-0.0005***	-0.0007***	-0.0005***	-0.0004***	-0.0005***
	(-24.74)	(-16.99)	(-25.89)	(-17.86)	(-5.94)	(-4.97)
Education dummies: No education is the reference						
Primary level	0.0113	0.0828***	0.1229***	0.1823***	-0.1028**	-0.0265
	(0.57)	(4.19)	(6.06)	(9.08)	(-2.18)	(-0.49)
Lower secondary	0.1931***	0.3190***	0.3458***	0.4409***	-0.1588***	-0.0346
	(9.12)	(14.74)	(16.16)	(20.27)	(-3.03)	(-0.53)
Upper Secondary	0.4798***	0.5767***	0.6284***	0.6851***	-0.0691	0.4823***
	(18.46)	(24.15)	(24.69)	(29.01)	(-0.58)	(3.59)
College & above	0.6817***	0.6938***	0.8238***	0.8010***	0.2943***	0.6348***
	(30.63)	(30.51)	(37.06)	(35.37)	(3.53)	(6.03)
child06	0.0080	-0.0375***	0.0030	-0.0434***	0.0230	-0.0185
	(1.49)	(-6.04)	(0.58)	(-7.39)	(1.47)	(-0.76)
child612	0.0028	-0.0242***	-0.0029	-0.0327***	0.0333**	0.0259
	(0.55)	(-4.28)	(-0.59)	(-6.06)	(2.41)	(1.31)
child1215	-0.0005	-0.0300***	-0.0042	-0.0255***	0.0366*	-0.0749***
	(-0.08)	(-3.91)	(-0.62)	(-3.42)	(1.92)	(-2.83)
Ethnic	-0.1216***	-0.0559***	-0.0719***	-0.0514***	-0.2486***	-0.0694*
	(-11.60)	(-4.83)	(-7.11)	(-4.67)	(-9.18)	(-1.73)
married	-0.0138	0.0088	0.0206	0.0212**	-0.1105**	-0.0864**
	(-0.93)	(0.85)	(1.48)	(2.17)	(-2.34)	(-2.48)
migrant	0.0539***	0.0174***	0.0633***	0.0237***	0.0225	0.0225
	(9.30)	(2.94)	(11.29)	(4.15)	(1.10)	(0.87)
Urban	0.1493***	0.0859***	0.1354***	0.0824***	0.0869***	0.0632**
	(21.00)	(10.83)	(20.31)	(11.11)	(4.55)	(2.33)
Sector dummies: agriculture is the reference						
Industry	0.3714***	0.2498***	0.1040***	0.0198	0.6813***	0.5158***
	(41.83)	(22.08)	(10.43)	(1.64)	(32.63)	(16.52)
Services	0.2674***	0.2029***	0.0312***	0.0092	0.4153***	0.2851***

	(27.19)	(16.63)	(2.95)	(0.72)	(15.43)	(7.76)
Region dummies: Hanoi is the reference						
region1	-0.1906***	-0.1414***	-0.1433***	-0.1058***	-0.2060***	-0.2693***
	(-12.88)	(-9.07)	(-10.59)	(-7.42)	(-4.34)	(-4.05)
region2	-0.1522***	-0.1899***	-0.1528***	-0.1819***	-0.0600	-0.0899
	(-9.98)	(-11.73)	(-11.01)	(-12.26)	(-1.26)	(-1.39)
region3	-0.1464***	-0.2128***	-0.1561***	-0.2065***	-0.0310	-0.1450**
	(-10.19)	(-13.65)	(-11.88)	(-14.40)	(-0.69)	(-2.34)
region4	-0.1380***	-0.1208***	-0.0923***	-0.1110***	-0.2259***	-0.1216*
	(-8.13)	(-6.79)	(-5.88)	(-6.74)	(-4.28)	(-1.74)
region5	0.1111***	0.0878***	0.0663***	0.0457***	0.1942***	0.1829***
	(7.03)	(5.28)	(4.60)	(2.99)	(3.87)	(2.77)
region6	-0.1650***	-0.2143***	-0.1694***	-0.2169***	-0.1062**	-0.1744***
	(-10.92)	(-12.89)	(-12.21)	(-14.08)	(-2.28)	(-2.77)
region8	0.1605***	0.1518***	0.1314***	0.1325***	0.2787***	0.2407***
	(9.44)	(8.17)	(8.55)	(7.84)	(4.87)	(3.03)
Types of ownership dummies: domestic private sector is the reference						
Public	0.1490***	0.3420***	0.0491***	0.2582***	-0.0871**	0.3061***
	(18.03)	(35.97)	(6.24)	(28.52)	(-2.31)	(6.13)
FDI	0.0937***	0.2390***	0.0856***	0.2338***	0.1611	0.2892***
	(5.54)	(17.87)	(5.48)	(18.81)	(1.55)	(3.55)
Constant	1.3204***	1.3749***	1.4106***	1.4736***	1.8480***	1.4014***
	(23.02)	(24.08)	(25.93)	(27.22)	(12.00)	(7.01)
Observations	50,073	36,842	43,623	33,409	6,450	3,433
Number of id	30,128	22,212	27,452	20,620	6,014	3,256
R-squared	0.357	0.418	0.335	0.398	0.306	0.221

z-statistics in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Source: Vietnam Labour Force Survey, 2012

Table A3: Returns to education by random effects estimation - Ownership and sector

VARIABLES	(1) FDI	(2) Public	(3) Private	(4) Agriculture	(5) Industry	(6) Services
Age	0.0158***	0.0970***	0.0241***	0.0073	0.0247***	0.0851***
	(2.64)	(32.95)	(10.53)	(1.41)	(11.32)	(31.73)
age2	-0.0002**	-0.0011***	-0.0004***	-0.0002***	-0.0003***	-0.0010***
	(-2.25)	(-33.37)	(-13.89)	(-3.28)	(-12.65)	(-33.21)
Education dummies: No education is the reference						
Primary level	0.0840	0.1939***	-0.0140	-0.0698***	0.1638***	0.0978**
	(1.31)	(3.13)	(-0.95)	(-3.06)	(8.16)	(2.27)
Lower secondary	0.2183***	0.5670***	0.0736***	-0.1596***	0.3170***	0.4240***
	(3.32)	(9.13)	(4.42)	(-4.78)	(15.03)	(9.76)
Upper Secondary	0.4137***	0.8349***	0.3406***	0.2165*	0.5278***	0.7143***
	(5.18)	(13.32)	(11.02)	(1.85)	(17.93)	(16.12)
College & above	0.8661***	0.9587***	0.6637***	0.4393***	0.8005***	0.8378***
	(12.13)	(15.39)	(31.09)	(7.32)	(33.67)	(19.18)
female	-0.1380***	-0.0271***	-0.2665***	-0.1636***	-0.2588***	-0.0924***
	(-9.35)	(-3.62)	(-38.79)	(-10.98)	(-40.61)	(-12.73)
child06	-0.0026	-0.0164**	0.0091*	-0.0051	0.0008	-0.0111*
	(-0.21)	(-2.52)	(1.69)	(-0.43)	(0.16)	(-1.83)
child612	0.0093	-0.0314***	0.0059	0.0021	0.0021	-0.0243***
	(0.79)	(-4.99)	(1.20)	(0.20)	(0.44)	(-4.08)
child1215	-0.0115	-0.0111	-0.0137**	0.0046	-0.0144**	-0.0136*
	(-0.61)	(-1.31)	(-2.08)	(0.34)	(-2.20)	(-1.69)
ethnic	0.0280	0.0077	-0.1780***	-0.1813***	-0.0907***	-0.0168
	(0.76)	(0.69)	(-16.19)	(-9.38)	(-7.38)	(-1.50)
married	-0.0139	0.0095	-0.0182	-0.1115***	0.0092	0.0166
	(-0.67)	(0.66)	(-1.63)	(-4.33)	(0.87)	(1.30)
migrant	0.0359**	0.0537***	0.0148**	-0.0174	0.0374***	0.0456***
	(2.57)	(9.04)	(2.44)	(-1.21)	(6.50)	(7.98)
urban	0.0939***	0.1617***	0.1046***	0.0917***	0.0975***	0.1502***
	(6.51)	(18.33)	(15.01)	(5.73)	(15.40)	(17.94)
Sector dummies: agriculture is the reference						
Industry	-0.0114	-0.1467***	0.3932***			

	(-0.13)	(-7.64)	(51.11)			
Services	0.0832	-0.3050***	0.3638***			
	(0.91)	(-17.71)	(38.91)			

Region dummies: Hanoi is the reference

region1	-0.3207***	-0.0693***	-0.2954***	-0.1398**	-0.1609***	-0.1243***
	(-7.48)	(-4.73)	(-18.21)	(-2.05)	(-11.87)	(-8.58)
region2	-0.2512***	-0.1635***	-0.1625***	0.1262*	-0.0921***	-0.2290***
	(-6.64)	(-10.04)	(-10.40)	(1.78)	(-7.21)	(-14.34)
region3	-0.2373***	-0.1513***	-0.1685***	0.2997***	-0.1401***	-0.2081***
	(-5.44)	(-9.85)	(-11.31)	(4.54)	(-10.88)	(-14.13)
region4	-0.3717***	-0.0881***	-0.1509***	0.1874***	-0.0322*	-0.1223***
	(-5.33)	(-5.20)	(-8.26)	(2.78)	(-1.77)	(-7.28)
region5	-0.0775**	0.0717***	0.1065***	0.7212***	0.0877***	-0.0899***
	(-2.16)	(3.98)	(6.60)	(10.85)	(6.38)	(-5.18)
region6	-0.1634***	-0.1083***	-0.2050***	0.2853***	-0.1593***	-0.2104***
	(-3.54)	(-6.33)	(-13.28)	(4.31)	(-11.37)	(-13.39)
region8	-0.0206	0.1381***	0.1656***	0.7161***	0.2070***	0.0839***
	(-0.56)	(6.43)	(9.93)	(8.34)	(14.52)	(4.65)

Types of ownership dummies: domestic private sector is the reference

Public				0.7374***	0.2636***	0.0673***
				(32.41)	(29.46)	(7.63)
FDI				0.7121***	0.1692***	0.2584***
				(4.70)	(18.03)	(8.34)
Constant	2.5880***	0.6183***	2.0513***	2.1147***	2.0936***	0.7792***
	(16.71)	(6.85)	(38.90)	(15.83)	(41.99)	(10.74)
Observations	4,575	38,706	43,634	12,086	31,918	42,911
Number of id	2,800	22,361	28,524	8,733	20,128	25,340
R-squared	0.338	0.288	0.312	0.322	0.329	0.327

z-statistics in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Source: Vietnam Labour Force Survey, 2012