OSIPP Discussion Paper: DP-2013-E-001

Is a gender gap in net school enrollment a reflection of the gender wage gap in the labor market? Evidence using household data from Vietnam

January 16, 2013

Tien Manh Vu†

Ph.D Candidate, Osaka School of International Public Policy,

Osaka University

[Key words] School dropouts・Returns to schooling・Wage・Gender gap・Vietnam
[JEL Codes] I24, I25, J31

[Abstract]
The paper estimates both the gender gap in wage and net schooling enrollment from Vietnam household data. The results imply a reflection of gender wage gap in the labor market in hazard of school withdrawals. Generally, males have higher incentive to terminate their schooling to join the labor force. Males would have 43.8 percent higher in participating the labor market and gain 18.4 percent of wage per hour higher than females. Also, we observe 16–44.4 percent lower in probability for males to enroll in school, especially, the school withdrawal rate accelerates at higher speed after the age of primary school. Meanwhile, females would have an incentive to complete junior, senior high school and 3–year college thanks to higher speed gain in wage. Besides, family having a combination of a household head working for a state–owned firm and his spouse working as self–employed would best facilitate their co–residing children and grandchildren for more years of schooling. Finally, the current education subsidy and tuition fee reduction policy do minimal to reduce the hazard of school dropouts among beneficiaries.

†Contact: Osaka School of International Public Policy, Osaka University, 1-31 Machikaneyama, Toyonaka, Osaka 560-0043 Japan. Tel/Fax: +81-(0)6-6850-5656. E-mail: t-vu@osipp.osaka-u.ac.jp.

[Acknowledgments] The author acknowledges the useful suggestions of Hisakazu Matsushige, Miki Kohara, Masaru Sasaki, Hiroyuki Yamada, participants in the graduate seminars chaired by Tsunehiro Ootsuki, and by Charles Yuji Horioka of Osaka University, and Rafael Lalive of University of Lausanne. All remaining errors are ours.

1 Introduction

During the economy transition, Vietnam have enjoyed high gross domestic product growth rate for almost over 5 percent per annum from 1990–2008 (GSO 2010) with increasing number of "active" firms[1] from 42,288 in 2000 up to 155,771 in 2007 (GSO 2012). Such shift needs more skilled workers and skill‐based wage would have established. For example, higher income for workers in productions and services would be a measure for newly established firms to absorb labors. Thus the choice of schooling more than primary education would become more important to each individual's future income. These assumptions encourage us to verify and compare the corresponding changes in school enrollment by gender.

The main purpose of the study is to investigate the gender gap in both wage and net school enrollment using data from the 2008 Vietnam Household Living Standard Survey (VHLSS). Also, we are interested in the efficiency of policies on education and linkage between parents' (grandparents') wage to the schooling choices of the children. Unlike previous studies, we examine the wage in the labor market by using in the Heckman sample selection models with greatest combination of 33 occupational types and 88 industries where 101,306 individuals are working. Meanwhile, the school enrollment gender gap is analyzed in Cox hazard models and regarding the effect the family background as well as education policies. In addition, we link education premium with schooling choice decision by using some control variables (which are highly valued in the labor market) in the hazard models and comparing the marginal effects between returns to education and hazard of dropouts.

We obtain four major findings. First, males are more likely to involve in salary work, at about 43.8 percent higher in probability and benefit 18.4 percent higher wage than females. However, females have a relatively increment incentive of 4.51‐9.49 percent for an additional degree from junior high school up to 3‐year college. Second, males are facing higher school dropout hazard and the higher acceleration to dropout than females from age 12‐22. Third, we find a combination of a household head working for a state‐owned firm and his spouse working as self‐employed would reduce the hazard of school dropout to the lowest. Finally, although the tuition fee reduction policy is widespread to almost 40 percent of individuals aged 7‐22, the policy has minimal connection to an additional year in schooling. Similarly, the education subsidy does not create any statistically significant gain for net school enrollment.

The remainder of the paper is organized as follows. Section 2 presents previous advances in estimating the gender gap in wage and school dropouts as well as theory bases for our hypothesis. Section 3 introduces the data used and Section 4 describes our empirical methods. Section 5 details our findings. Section 6 provides our conclusions.

[1] "Active" means the firms actually have employee to do business and create some business cost during the year.

2 Background to gender gap in wage and schooling choice

2.1 Advances in estimating gender wage gap and school choices

The returns of schooling have been examined for several decades since Mincer (1958). The returns of schooling would vary among population with observable factors such as school quality and education of the parents (Card 1999). Heckman, Lochner and Todd (2008) estimate the marginal internal rates of return for different schooling levels and find relatively larger returns to graduating from high school than to graduating from college given that both have been increasing overtime.

Schooling contains endogeneity problems. The common is the unobservable of "ability". Second, schooling would be the optimization behavior of individuals and families (Griliches 1977). Third, the aggregate log of earnings can be right skewed in contrast with original Roy model that assumes log skills are normally distributed (Heckman and Honoré 1990). Furthermore, we learn from Card (1999) that it would be difficult to decide whether the higher education cause higher earnings observed from better educated individuals or whether individuals with greater earning capacity chose to have more years of schooling. Thus we realize intercepting the schooling decision in the early life would be important because wage earning capacity is still potential.

The estimation technique is improved significantly by the development of theory and econometric methods to get an insight to unobservable characteristics, ability. Measurement errors can be justified by using potentially instrumental variables such as parental education (Willis and Rosen 1979), tuition fees, minimum school‑leaving age, college proximity (Card 2001) and test score (Card 1999), and using identical twins data (Ashenfelter and Krueger 1994) where genetic difference in endowments can be controlled.

There are some explanations for gender wage gap. The first reason could be the difference in gender specific factors. For example, women are likely to be involved in shorter and discontinuous work due to childbirth, child rearing and housework (Mincer and Polachek 1974) and time constraint between market work and house work (Becker 1985). Thus women are more likely to have less experience and/or less incentive to invest or be invested in on–job training and education. Similarly, women with different marital status and number of children devote different proportion of her lifetime to the world of work (Mincer and Polachek 1974). Second, the wage structure (distribution) could explain for the gap. Fortin and Lemieux (1998) indicate the decline in the gender wage gap during 1979–1991 was highest at the middle centile of the wage distribution. This is due to the distribution of wage function of females became less skewed to the right and had similar shape to that of males. Also, we acknowledge that the gender wage gap would be lower at the entry of the labor market, larger in the middle and the other end of exit (Tansel and Bodur 2012). Topel (1994) shows wage inequality differs by regions. On the other aspect, changing the nature and condition of work, i.e increasing computer use, would open more job opportunity to females (Weinberg 2000). There is also an unexplained gap. For example, Fortin (2005) finds the impact of soft factors such as greed, ambition, altruism, locus of control, gender role attitude, family values and income expectations would have impact on behaviors, wages and gender wage gap. Finally, the gender wage gap can be a purely gender discrimination. Goldin and Rouse (2000)

find the shift to "blind" auditions can explain 25 percent of the increase in the percentage of female members in the orchestras from 1970–1996. Because music directors in charge of hiring new musicians, publicly disclosed their belief that females had lower musical talent.

Literature reports several ways to decompose the gap. Detected gender differences in wage consist of differences in the endowments of the sampled individuals, in coefficients and unexplained part, probably due to discrimination. Works by Oaxaca (1973) and Juhn, Murphy and Pierce (1991) are of the extent and later widely applied in other studies. However, the decomposition considers only the mean wage difference in the two latter parts. Thus quantile regressions further examine whether the gender wage gap is larger (smaller) in either ends of wage distribution (Buchinsky 1994, Gardeazabal and Ugidos 2005, Tansel and Bodur 2012).

Some studies examine the school choice focusing in college degrees in developed countries and primary schools in developing countries. For example, Willis and Rosen (1979) observe the expected lifetime earnings would influence the decision to attend college. Brown and Corcoran (1997) find the gender difference in choosing schooling content relates to gender wage gap among college graduates. Meanwhile, increasing returns to education would raise (primary) school enrollment (Foster and Rosenzweig 1996). Similarly, increasing wealth would induce parents invest more for children's education in Vietnam (Glewwe and Jacoby 2004). However, literature reports some time lags or even mismatches at the time that individual left school. Card (2001) indicates different individuals would finish their schooling at the time when the marginal return to last year of schooling would either above or under the average marginal return of the population. However, using the National Longitudinal Survey of Youth (NLSY) 1979, Cunha and Heckman (2007) find 50 percent of the ex–post variance in returns to college is predictable at the time the individual making the choice of schooling. More specifically, the study predicts that 13.86 percent of high school graduates would rather be college graduates and 17.15 percent of college graduates would have stopped their schooling with a high school diploma. Besides, Glewwe and Kremer (2006) suggest that the elasticity of demand for schooling is more likely higher for females than males in developing countries. Thus policies and programs on education may increase enrollment rates of females without having the intended target.

2.2 Theory base for our hypothesis

In decision theory, individuals make decision on schooling choice by maximizing their utilities. We assume that individuals are able to evaluate their learning ability. We also assume that their learning ability is unlimited given that they are provided with enough resources to learn. Hence, individuals consider the expected present value of an additional year in schooling, k, versus the expected forgone income in $T - t - k$ years if they work and the cost of learning, and other unobservable utilities such as preference on learning, psychic cost, job searching cost.

$$\text{Education choice}_t = \text{f(abilities, family background, geographic differences, } \dots \text{, and expected wage value)} \quad (1)$$

More specifically, individuals make decision of schooling choice at time t

$$\text{Education choice}(t|\text{ability}) = \text{Max}_k \{E_t(V_{T-t-k}(.)) - C_k - E_t(W(k, edu_t))\} + e \quad (2)$$

in which, $E_t(V_{T-t-k}(.))$ is sum of expected future income after obtained the degree of k years of schooling. $C_k + E_t(W(k, edu_t))$ are cost to obtain the degree and expected forgone income at current education level within k additional year of schooling. e is other unobservable (unmeasurable) utilities such as: abilities, preference, psychic cost (job searching cost), etc.

We suppose that all individuals can observe the wage function and/or distribution, e.g from the wages of their relatives and acquaintances. The present value of current expected log wage (if the individual is working) is

$$\text{Log wage}_t = \frac{\log(W(k, edu_t))}{hours*(1+k)^s} , \quad (3)$$

where s is the schooling premium. Therefore, expected forgone wage at current education level can explain for the education choice. The wage gender gap, in turn, as a part of wage function, can also explain for the education choice.

3 Data

The data in use is cross section data, common known as Vietnam Household Living Standard Survey 2008. The sample size is 45,945 households with 289,948 individuals. The sampling method is two–stage stratified by using 3,045 sample units taken from master sample of population census. Two parts of the survey are containing necessary information about income and education.

For income analysis, our definition of labor force and definition of working are different from the concept from the survey. The labor force in our concept consists of two types of people, working and non–working individuals. Working individuals are defined as people who have wage (salary) in the year 2007 for their main job. By the definition of the survey, main job is an activity consuming most time among other jobs (economic activities). In Vietnam, an individual can have more than a single job and economic activities to generate income. For example, an individual can work as civil servant during the day and as self–employed for family business at night, such as producing ceramic crafts. VHLSS 2008 collects detailed information of first main job[2], the second job and so on. Non–working individuals are those who do not have wage (salaries) for their first main job. They are seeking for such kind of job or doing some other activities as substitutions for the main job. Besides, we further restrict the labor force by age and possibility to (re)enter to the labor force. Individuals in the labor force are at the age permitted by Vietnamese Laws of Labor. They have neither retired nor reached to the age of retirement by the laws. In addition, they are neither enrolling in any kind of school nor expecting for school entrance exams. Labor force does not contain individuals who are disabilities, too weak to work, facing serious illness and taking early retirements. Descriptions of selected sample for income analysis are as in Table 1.

[2] Main job is defined in VHLSS as earnings required the highest number of time devoted to. If an individual has two equal time spending on jobs, highest earnings job is recorded as main job.

[INSERT TABLE 1 HERE]

From the sample, we can observe a "labor force" relatively young with an average of 36.33 years of age and females occupy 49.25 percent. In addition, paid work as main job (most time consuming and highest earning job) attracts only 33 percent of the total labor force. 43.97 percent of individuals are self-employed, 72.59 percent of which is in agriculture, forestry and fishery sectors. Therefore, we argue that sample selection would be the most serious problem in estimating the returns to schooling. Also, we notice that the distribution of wage distribution of both males and females are right-skewed although that of males is less than that of females as seen in Graph 1. The gender gap would be minimal in the far right tail whereas the gender gap would be significantly different in the left and the right middle of the distribution.

[INSERT GRAPH 1 HERE]

For education analysis, we define net school enrollment as individuals who are enrolling in any kind of school at the year 2008. However, the individuals are 7–22 years of age and do not repeat any grade[3] since they have entered to primary school. This is important as suggested by Glewwe and Kremer (2006)[4]. Besides, we define school leavers as individuals aged 7–22 and have a record of school attendance[5]. The sum of school enrollments and school leavers confound the total individuals who have been made decision on schooling. In order to further drill on the time of making decision, we construct two panels. Panel 1 is as what we defined. Panel 2 further restricts to individuals who were enrolling in a school in the year 2007. Therefore, Panel 1 deals better with accumulation of schooling choice made by individuals of interest in all past years while Panel 2 facilitates to analyze the school decision made in 2007. Panel 2 also enables us to integrate some educational policy variables. Descriptive statistics for the two panels are shown in Table 2 and 3.

[INSERT TABLE 2 AND 3 HERE]

4 Empirical methods and specifications

4.1 Heckman sample selection with two steps for wage regression

We apply Heckman sample selection model with two steps (Heckman 1979) to estimate the gender wage gap as we are unable to observe wage of approximately 68 percent of total individuals who are engaged (ready to be engaged) with some business activities[6].

We acknowledge that the model is not able to decompose the gender gap into

[3] Appendix 1 shows descriptions of Vietnam education system.

[4] The study indicates principals and teachers have incentives to exaggerate the number of students enrolled in developing countries. Thus the definition of net enrollment rate is more appropriate. Net enrollment rate is the number of children enrolled in a particular level of schooling who are of the age associated with that level of schooling divided by all children of the age associated with that level of schooling (Glewwe and Kremer 2006).

[5] We acknowledge that 1,327 individuals (2.5 percent) in Panel 1 aged 7 or above have never entered to any school.

[6] Glewwe and Kremer (2006) record this problem is serious in developing countries.

difference of endowments, of coefficients and of unobservable factors. Also, the model does not consider the interaction between the coefficients of individual characteristics with the gender. In addition, the model does not divide data into different quantiles of wage and consider the gender wage gap in each quantile. However, we would argue that sample selection would be the most serious problem in detecting the gender wage gap in Vietnam. Besides, the general mean of gender wage gap is of our interest because it is easiest to be observed by any individual. Thus the general mean of gender wage gap is more likely to be the signal for each individual to align their years of schooling at their early life. The quantile gender wage gap would be more precise and more important for individuals at their mid‑career. Thus we try to reduce the effect of occupational types and industries' differences, which are more likely to lead to different gender wage gap in different part of the wage distribution, by introducing corresponding control variables.

We assume that the selection can be predicted from the individual characteristics X_i such as sex, marital status (Das et al. 2003), age, work experience, from the place of residence such as rural or urban, employment opportunities (economic regions can be a proxy), and highest educational attainment. In the selection stage,

$$z'_i = select_i = \begin{cases} 1, \ if z_i = X_i\beta + \varepsilon \geq z^* \\ 0, \ if \ z_i < z^* \end{cases}. \quad (4)$$

In the outcome stage,

$$y_i = \log wage_i = X'_i\beta' + \epsilon. \quad (5)$$

The wage estimate is Mincer regression and can be explained by X'_i characteristics such as sex, age, working experience, rural or urban area of residence, economic regions, highest diploma obtained, highest vocational degree obtained, characteristics of the enterprise such as types of ownership and industries.

4.2 Hazard model for the probability of leaving school

Meanwhile, we deploy a hazard model to estimate the probability of withdrawing (dropout) from school for individuals aged 7–22. We omit all individuals who are attending in schools but their ages are more than 2 years difference to the age of an individual starting primary school at 7 years of age and continuously adding one grade annually. The period of study is the time since an individual enters primary school at 7 years of age until either the individual leaves the school or the survey ends. At the age 7, all individuals are considered to be at school as primary school is compulsory in Vietnam. The individuals are also at risk of dropout every year after the age 7. If an individual does not have a status of "at school" in 2008, the individual is considered as already left school (termination). If the individual is "at school", the individual is regarded as right censored data. It is reasonable to assume that all censored data will fail or finish their schooling sometime after 2008.

Suppose that the year of schooling, started from 7 years of age when an individual enters primary school and finishes when the individual terminates

schooling, follows $F(t)$ [7]. The survival function, indicating the probability that an individual continue for higher school grade at time t, is then $S(t) = 1 - F(t) = P(T > t)$. $S(t) = 1$ at $t = 0$. The hazard function, $h(t)$, is the instantaneous rate of failure or the rate of terminating school in such a short period of time:

$$h(t) = \lim_{\Delta t \to 0} \frac{P(t+\Delta t > T > t | T > t)}{\Delta t} = \frac{f(t)}{S(t)}. \quad (6)$$

The cumulative hazard function is

$$H(t) = \int_0^t h(u).du = \int_0^t \frac{f(u)}{S(u)} = -\ln(S(t)). \quad (7)$$

In general, the likelihood function of an individual experiencing school termination is

$$\mathcal{L} = S(t).h(t) \quad (8)$$

In case of right censored data

$$\mathcal{L} = S(t) \quad (9)$$

Thus the likelihood function of both censored and uncensored individuals can be rewritten as

$$\mathcal{L} = \prod_{i=1}^{n} [\{S(t_i)\}^{d_i} . \{S(t_i).h(t_i)\}^{1-d_i}] \quad (10)$$

4.3 Specifications

4.3.1 Dependent variables

In wage regression, the dependent variable is logarithm of wage per hour in 12 month prior the survey. The wage per hour is the ratio of total wage by total working hour. The total wage is a sum of the total wage, bonus and any other income related to wage. Meanwhile, the total working hour is a product of sum of total hour working per day, total days of working in a month, and total months of working in 12 months prior to the survey.

In school dropout analysis, the dependent variable is the years of education. If the individual has left school, years of education is the highest school grade ever be completed. If the individual is enrolling in a school by the time of survey, it is the current grade minus 1 by the definition of VHLSS.

4.3.2 Independent variables

Independent variables in wage analysis include individuals characteristics such as sex, marital status, age, working experience, highest obtained academic degree, highest obtained professional degree, residing area (rural/urban and economy regions), ownership of the enterprises (classified by GSO), 88 industries where the employers are mainly involving, and 33 occupation types (classified by GSO).

In school dropout estimates, independent variables consist of individuals and family characteristics including sex, household income (household income adjusted by root squared of number people in the household), whether the individual is the only son and/or the grandchild of the household head and residing area (rural/urban and economy regions). Also, we add information about the household head and his

[7] We simplify all functional forms without displaying covariates. For example, $h(t)$ is used instead of $h(t; X)$, where X denotes independent variables.

spouse[8], including education, their employers' type if they are working. This would be a reasonable assumption as the difference in wage among household head and his spouse, if any, would be the easiest to be observed. Besides, the information could help to control a possible interference of household heads (spouses) to the schooling choice of children within the household, given the heads (spouses) acknowledges of their wage premium. Also, working in different employer types may result in different time schedule which might influence the time for rearing or support their children (grand children). However, since there are 5,498 (1,327) individuals missing information about their household head (head's spouse), we analyze the information in a separate estimation. Also, in a separate data, the Panel 2, we add three dummies for the status of education policy recipients such as tuition fee reduction, scholarship, education subsidies. Ratio of tuition fee over the total expenditure on individuals' education is included. The information only appears among individuals who experienced school enrollment in 2007. Detailed explanation for both dependent and independent variables can be seen in Table 2, 3 and 4.

[INSERT TABLE 4 HERE]

In addition, we introduce an interaction term between sex and the variable of interests. In wage function, it is the highest degree obtained by the individuals. If the interaction term is significant, the coefficient of the interaction term can explain for the difference in the incentive to complete the highest degree among the sex of interest. In a hazard model for school dropouts, the corresponding coefficient would show the difference in effort to add an additional year in schooling among the sex of interest. As the magnitude of the interaction effect in non-linear models is not equivalent to the marginal effect of the interaction term and the its statistical significance is not calculated by standard software (Ai and Norton 2003). Thus, we apply the tips by Buis (2007) to calculate and interpret the interaction term.

4.4 Marginal effect over sex and the variable of interest

There are three main reasons for us to conduct marginal effect over sex and the variable of interest. The first is that Heckman et al. (2006) present evidence that the marginal return does not equal to the average return which is often estimated from Mincer regression. The second is lying to the fact of the coefficient of the interaction term in non-linear models is not equivalent to the marginal effect. Besides, we acknowledge that the marginal effect to complete a degree would be higher in the final year of degree completion (sheepskin effects). Heckman, Lochner and Todd (2008) indicate that the levels and trends in rates of return estimated from Mincer model are misleading for many schooling levels.

5 Results

5.1 Differences in labor participation rate and gender gap in wage

Estimates of the probability of participation would capture the general image of the labor market of Vietnam, as shown in Table 5. People living in different economic

[8] In fact, we convert all female household head as household head spouse and male household head spouse as household head for the selected data.

regions and/or rural (urban) area also experience different labor participation rate. Highest of these rates are Southeast and Red River Delta (as the base). Probably, the two regions embrace two largest cities in Vietnam, Ho Chi Minh City and Hanoi where most firms are located. In general, estimates show individuals with higher level of education tend to have higher participation rate, which agrees with Becker (1962) or they are self-selected to the labor market. Meanwhile, FDI firms, state–owned firms and private-owned firms would be attractive to employees, at least novice, for paying 29.6, 18.9 and 15.3 percent higher wage per hour than collectives and family owned business, respectively. However, the percentage of individuals having more than 12 years of schooling among state-owned firms in our selected sample is 64.35 whereas the numbers are 27.05, 28.47 and 3.38 in FDI firms, private firms and family business, respectively. Thus state-owned firms' employees would have a different pay scale which may put heavy weights on education degrees. In addition, the employees would have added up their degrees while working as a response to that scale. Similarly, a 3-year vocational course graduates can earn 11.5 percent higher than the one without the degree. Besides, while coefficients of region dummies in the first step illustrate the probability to participate in the labor market, the corresponding coefficients in the wage regression show the regional wage difference. Though Red River Delta would be where proportion of wage earners is the country second highest rate, the region is also the country second lowest wage paid per hour, holding other else constant. Besides, the return to education would be increasing if individuals hold higher degree up to as equal as master degree and then slightly decreases. Especially, upper secondary graduates would receive 15.4 percent higher wage than the base while the return would jump to 33.7 if individuals hold some college degrees.

[INSERT TABLE 5 HERE]

Females are less likely to participate in labor market, with 43.8 percent lower in probability and suffer a wage of 18.4 percent lower (as seen in Table 5). Especially, the interaction term between gender and the highest degree show an interesting feature. Although experiencing from the wage gap of 12.4 percent lower than males, comparing among females, females would have additional wage gains by higher education. As shown in Table 5, the interaction terms between female and degree of junior, senior high school, 3–year college, and 4 or 5 year university can explain for a corresponding 4.51, 6.19, 9.49 and 6.0 percent wage gains. The increment in wage gains would be an incentive for females to accomplish the four year degrees. Nevertheless, the incentive would be minimal in primary school as the corresponding coefficient is statistically insignificant. It is very likely that compulsory education during primary school with no tuition fee fill the gap if any.

5.2 Differences in probability to leave school and the gender gap

In general, individuals at age groups 12-15, 16-18, 19-22 would be 100, 243.8 and 419.6 percent more likely to leave school than the age group 7-11 as seen in the Panel 1, Table 6. There are several possible reasons. First, the age group 7-11 is under compulsory education. Second, individuals have different ability, and people with higher ability are scarce than the average. Third, there would be lower accessibility to higher level of education, such as under-developed higher education services. Fourth, individuals may leave school for economic reason, such as for working. However, in

Panel 2 consisting of relatively higher ability people, the corresponding coefficients show a decrease from 85.4 to 98.31 percent in dropout rate in older age groups. Sheepskin effect, the phenomenon of highest return to the last grade of one education level, can also be an explanation. As individuals try to get the degree, they would not dropout in the middle of one school level. This agrees with our interpretation in Panel 1.

[INSERT TABLE 6 HERE]

Besides, differences in school dropouts among economic regions are significant. In Panel 1 shown in Table 6, individuals in North Central Coast would be 5.3 percent less likely to dropout. Interestingly, the region is often considered as poorest. Mekong River Delta region would have the highest school dropout rates in the country, with 148.2 percent probability higher than Red River Delta. In Panel 2, differences among economic regions are less serious but the evidence of country's highest dropout rate in Mekong River Delta is confirmed with 15.8 percent difference. Individuals in North East are 27.2 percent higher in probability to dropout school. In contrast, South East would have a lower dropout probability.

Family background would help to control different aspect to the probability to dropout as decision of school choice would be an interaction among individuals, their family (Griliches 1977) and the society. We argue for the role of society such as educational policies and labor market settings. We prefer to add such information of family background to trade off with bias selection (4,475 individuals are left behind). Apart from the core variables which have consistent effect in the previous analysis, information of family background would explain some interesting facts. In Table 6, analysis 2, higher educated household head and his spouse would lead to 7.9-8.7 lower probability to leave school earlier. Interestingly, there is a different attitude of household head and spouse to their children and/or grandchildren. We once observed the highest returns to schooling among individuals working for FDI, state-owned, and private firms, respectively. However, the school dropout hazard is highest among families having the head working for FDI firms (62.5 percent higher) and lowest among a combination of the head working for state-owned firms (with 13.3 percent lower) and the head spouse is self-employed (with 10 percent lower). Where the head is in FDI firms (in private firms), their children and grandchildren are facing an excessive dropout hazard of 62.5 (30.5) percent as shown in Table 6 (7), analysis 2. It could be a reflection of wage inequality to the next generation's schooling choice either in an active or inactive manner. A possible explanation is that the difference can be a result of difference in time flexibility of the head (head spouse) available for child rearing. If education can be a proxy for skill-based pay which is more likely to applicable in non-state sector, the differences between effect of FDI (private) firms and that of family business are more likely the time rearing the children. For instance, if the head spouse is self-employed, she would gain more flexible working time and thus having more time to educate the next generation.

[INSERT TABLE 7 HERE]

Household income level would lower down the probability of dropout to 7.2-30.4 percent as shown in Table 6 and 7. In addition, if the individuals are living with

grandparents, they are more likely to be kept in schools with 18.1-51.8 percent increment of probability compared with the others living without grandparents. Notably, the difference between the individual who is the only co-residing son and the base is 18.1 percent higher as shown in Table 6, analysis 1. We once try to establish one-daughter family to compare single gender effect. Unfortunately, the sample does not contain any of such family. Probably, the son preference leads to a higher investment on the education of the only son in the family. However, in other analysis, the corresponding coefficients are statistically insignificant.

Panel 2 allows us to have an insight to the effect of educational policies on the dropout rates (net enrollment rates). A tuition fee reduction would contribute minimal to the dropout rates of the individuals despite that 41.9 percent of the individuals are the beneficiary of this policy (see Table 4). Similarly, education subsidy receivers would have the same probability as others to dropout. In contrast, scholarship receivers would benefit 29.2-30 percent higher in probability to continue their schooling. Perhaps, there would be two kinds of scholarship receivers: those who have excellent academic records and those who have some economic disadvantages. Thus if individuals are only the former, the coefficients show their ability to continue higher education level. If they are the later or both, it is difficult to have policy implication from the result. In addition, the ratio of tuition fee payment over the total expenditure on education would have an interesting implication for policies on education. If the ratio increases one unit, it explains for 71.6 percent increase in probability to be in schools in the following year. In Vietnam, public education dominates and the tuition fee is set by the authorities. Thus policies to decrease or increase tuition fee cannot determine themselves the probability to dropout rates. Instead, the sensitive to education expense and the other component of education expense than tuition fee should be considered. On the other hand, if tuition fee is almost all education expense, individuals are more likely to be in school, otherwise the inverse applies.

The gender could explain for 16-44.4 percent higher in probability for females to stay in school. In other words, females are less likely to terminate their schooling than males. However, among females, those aged 16-18 are 27.2-31.5 percent more likely to dropout than other females aged 19-22 in Panel 2. We will further present the difference in the following sub–section.

5.3 Gender gap in wage and school withdrawals

In wage regression analysis, we have estimated the interaction term between gender and the degree obtained, which could be an incentive for females to stay longer in school. The result suggests that the incentive would explain for the lower school dropout rate of females. More specifically, females would have incentive to complete junior, senior high school, 3-year college and 4 or 5 year university than males. In hazard model for school dropouts, we observe an inverse corresponding gender dropout gap of 27.2-31.5 percent.

[INSERT TABLE 8 HERE]

We further analyze the marginal effect over gender and one variable of

interest in both wage and school dropout analysis. In general, males often have higher marginal effect over wage per schooling level as shown in Table 8. The difference would explain for higher (lower) relative school withdrawal (dropout) rates among lower (higher) ability individuals. In addition, the corresponding wage gap (the speed of gains) would still be the incentive for females staying longer in junior, senior high school and 3–year college. In Table 8, the marginal effect is higher for females when moving from primary school to junior high school (1.801 compared with 1.052), from junior high school to senior high school (2.945 and 3.015), and form senior high school to 3‑year college (53.466 and 164.315). Meanwhile, as seen in Table 9, the marginal effect of hazard rate (also, the speed lost) over sex and age group is higher for males in all age groups. These results confirm and validate our previous analysis. Females have incentive of future return to education to move (complete) up to junior, senior high school and 3‑year college. In contrast, males are more likely to join the labor force than females.

[INSERT TABLE 9 HERE]

Unfortunately, we are not able to divide the age group 19‑22 to different learning routes from the data, for instance, first year in 3‑year college is treated as first year in 4 or 5 year university. Thus the marginal effect of hazard rate over sex and age group when an individual moves from age 16‑18 to age 19‑22 is the total effect of different routes. Therefore, we are unable to provide explanation for the differences in hazard marginal effect from the two age groups. In addition, as we limit the age of individuals in the data of school dropout analysis to 22, we are also unable to compare the later progression to master and doctor degree.

6 Conclusions

The study has examined the gender gap in wage among wage earners aged 15–60 (15–55) for males (females) and school withdrawal (dropouts) hazard among individuals aged 7–22. There would be a corresponding higher school withdrawal (dropouts) rate among males while the schooling premium is generally higher for males than females. We also observe the wage incentives among women to complete junior, senior high school and 3–year college. The speed of withdrawal (dropouts) among males would be higher at the corresponding age groups. Besides, the younger generation is given a relatively higher education while the only co‑residing son of the household head would have a similar benefit. Meanwhile, the employer types, which show a significant difference in wage, of the household heads and his spouses, in turn, would have a significant influence to the hazard of school withdrawal (dropouts) in the (grand) children co–residing in the household. More specifically, a combination between a head working for a state–owned firm and a head spouse working as the self–employed would decrease the hazard of dropout from school of the (grand) children to the lowest.

We acknowledge some limitations in our study. First is the limitation of the cross section data in use. The schooling premium may differ or change overtime according to the difference in returns to skills in the labor market (Heckman et al. 1998). Similarly, the dropout in Panel 2 is just the figure of period 2007–2008 which

can be affected by the economic settings. However, to our best knowledge, there would be no panel data and/or longitude data from Vietnam to eradicate the problem. Second, the hazard model replies heavily from data selection and distributional assumptions. Third, there can be a problem of sample selection in hazard model as 1,327 individuals (2.5 percent) in Panel 1 have never entered to any school. Fourth, omitted variables are always problems to both estimates of returns to schooling and estimates of hazard dropouts. For example, our estimates do not consider the difference in accessibility to school among individuals, quality of school and teachers as suggested by Glewwe and Kremer (2006). However, we prefer this to be able to estimate some key factors from the labor market as a guide to choose number of years in schooling. In addition, we are unable to distinguish the difference in career choices in post high school and link with the difference in the industries and occupational types. Nevertheless, we believe future studies would overcome our limitations when data are adequate.

Reference

Ahn, H., & Powell, J. L. (1993). Semiparametric estimation of censored selection models with a nonparametric selection mechanism. *Journal of Econometrics, 58*(1–2), 3–29.

Ai, C., & Norton, E. C. (2003). Interaction terms in logit and probit models. *Economics Letters, 80*(1), 123–129.

Ashenfelter, O., & Krueger, A. (1994). Estimates of the economic return to schooling from a new sample of twins. *The American Economic Review, 84*(5), 1157–1173.

Becker, G. S. (1962). Investment in human capital: A theoretical analysis. *Journal of Political Economy, 70*(5), 9–49.

Becker, G. S. (1985). Human capital, effort, and the sexual division of labor. *Journal of Labor Economics, 3*(1), S33-S58.

Brown, C., & Corcoran, M. (1997). Sex‐based differences in school content and the male‐female wage gap. *Journal of Labor Economics, 15*(3), 431–465.

Buchinsky, M. (1994). Changes in the U.S. wage structure 1963–1987: Application of quantile regression. *Econometrica, 62*(2), 405–458.

Buis, M. L. (2010). Stata tip 87: Interpretation of interactions in nonlinear models. *Stata Journal, 10*(2), 305–308.

Card, D. (2001). Estimating the return to schooling: Progress on some persistent econometric Problems. *Econometrica, 69*(5), 1127–1160.

Card, D. E. (1999). The causal effect of education on earnings. In O. Ashenfelter, & D. Card (Eds.), *Handbook of Labor Economics* (Vol. 3, Part A, pp. 1801–1863): Elsevier.

Cunha, F., & Heckman, J. J. (2007). Identifying and estimating the distributions of ex post and ex ante returns to schooling. *Labour Economics, 14*(6), 870–893.

Fortin, N. (2005). Greed, altruism, and the gender wage gap. Working paper, University of British Columbia.

Fortin, N. M., & Lemieux, T. (1998). Rank regressions, wage distributions, and the gender gap. *The Journal of Human Resources, 33*(3), 610–643.

Foster, A. D., & Rosenzweig, M. R. (1996). Technical change and human–capital returns and investments: Evidence from the Green revolution. *The American Economic Review, 86*(4), 931–953.

Gardeazabal, J., & Ugidos, A. (2005). Gender wage discrimination at quantiles. *Journal of Population Economics, 18*, 165–179.

General Statistics Office of Vietnam (GSO) (2010). Gross domestic product at constant 1994 prices by economic sector. http://www.gso.gov.vn/default_en.aspx?tabid=468&idmid=3&ItemID=12106. Accessed 9 July 2012.

General Statistics Office of Vietnam (GSO). (2012). Enterprises. Accessed 14 Nov 2012. http://www.gso.gov.vn/default_en.aspx?tabid=479&idmid=5.

Glewwe, P. (2002). Schools and skills in developing countries: Education policies and socioeconomic outcomes. *Journal of Economic Literature, 40*(2), 436–482.

Glewwe, P., & Jacoby, H. G. (2004). Economic growth and the demand for education: Is there a wealth effect? *Journal of Development Economics, 74*, 33–51.

Glewwe, P., & Kremer, M. (2006). Chapter 16 Schools, teachers, and education outcomes in developing countries. In E. Hanushek, & F. Welch (Eds.), *Handbook of the Economics of Education* (Vol. Volume 2, pp. 945–1017): Elsevier.

Goldin, C., & Rouse, C. (2000). Orchestrating impartiality: The impact of "blind" auditions on female musicians. *American Economic Review, 90*(4), 715-741.

Griliches, Z. (1977). Estimating the returns to schooling: Some econometric problems. *Econometrica, 45*(1), 1–22.

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

Heckman, J. J., & Honoré, B. E. (1990). The empirical content of the Roy model. *Econometrica, 58*(5), 1121–1149.

Heckman, J. J., Lochner, L., & Taber, C. (1998). Explaining rising wage inequality:

Explorations with a dynamic general equilibrium model of labor earnings with heterogeneous agents. *Review of Economic Dynamics, 1*(1), 1–58.
Heckman, J. J., Lochner, L. J., & Todd, P. E. (2006). Chapter 7 Earnings functions, rates of return and treatment effects: The Mincer equation and beyond. In E. Hanushek, & F. Welch (Eds.), (Vol. 1, pp. 307–458).
Heckman, J. J., Lochner, L. J., & Todd, P. E. (2008). Earnings functions and rates of return. *Journal of Human Capital, 2*, 1–31.
Juhn, C., Murphy, K. M., & Pierce, B. (1991). Accounting for the slowdown in black–white wage convergence. *Workers and their wages: Changing patterns in the United States*, 107–143.
Mincer, J. (1958). Investment in human capital and personal income distribution. *Journal of Political Economy, 66*(4), 281–302.
Mincer, J., & Polachek, S. (1974). Family investments in human capital: Earnings of women. *Journal of Political Economy, 82*(2), S76–S108.
Oaxaca, R. (1973). Male–female wage differentials in urban labor markets. *International Economic Review, 14*(3), 693–709.
Tansel, A., & Bodur, F. B. (2012). Wage inequality and returns to education in Turkey: A quantile regression analysis. *Review of Development Economics, 16*, 107–121.
Topel, R. H. (1994). Regional labor markets and the determinants of wage inequality. *The American Economic Review, 84*(2), 17–22.
Weinberg, B. A. (2000). Computer use and the demand for female workers. *Industrial and Labor Relations Review, 53*(2), 290–308.
Willis, R. J., & Rosen, S. (1979). Education and self–selection. *Journal of Political Economy, 87*(5), S7–S36.

Appendix 1 General descriptions of education system in Vietnam

Since 1995, the Laws of Education in Vietnam have been changed two times, in 2005 and 2009. However, the main characteristics of grade remain. Individuals enter primary school at age 6 as a compulsory education. After 5 years in primary school, individuals can enter to lower secondary school which lasts for 4 years. Right after lower secondary school, individuals have two choices but both require an entrance examination. The first is to enter upper secondary school for 3 years. The second is to join professional high school for 4 years whose graduates are recognized as equivalent to upper secondary school graduates plus a recognized degree in professional education and training (the degree holders have different wage ranking regulated by laws for employees in public sectors). After successfully passing the exams to be recognized as an upper secondary graduate, an individual can join entrance exams to the university which last for 4 or 5 years or for 3 year college which has lower requirement and less competitive entrance exams. The entrance exams to universities and colleges are regulated by Ministry of Education and Training in cooperation with each academic agent. Individuals are encouraged to take entrance exams to universities, if fails, their exams result is still usable for the selection to 3 year college.

Graph 1 Kernel density estimations of log wage per hour by gender

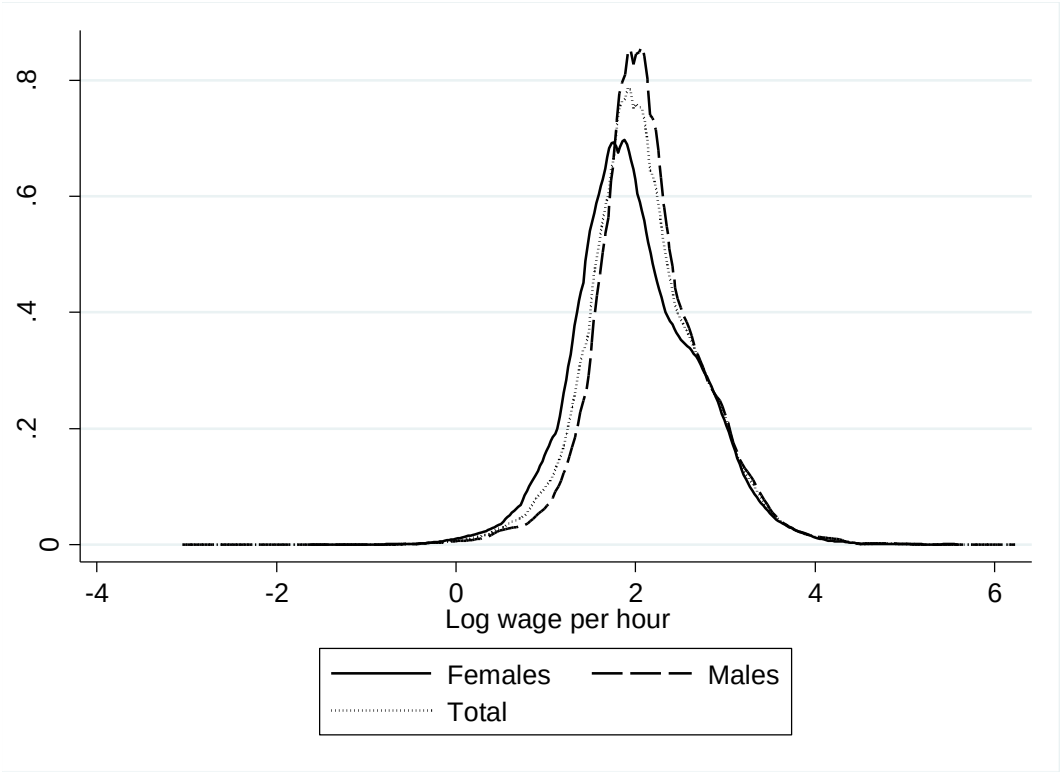

Table 1 Descriptive statistics on the labor force in VHLSS 2008

Category	Total	Males	Males (%)
Labor force	101,306	51,413	50.75
1. Working	33,425	20,513	61.37
2. Not working	67,881	30,900	45.52
2.1 Self–employed in agriculture, forestry and fishery sectors	32,336	14,817	45.82
2.2 Self–employed in other sectors than agriculture, forestry and fishery sectors	12,208	5,211	42.69
2.3 Do not have the first main job but have the second or third job	9,913	4,649	46.89
2.4 Do housework only	3,457	276	7.98
2.5 Cannot find a job	770	487	63.25
2.6 For some other reasons, they cannot work	986	533	46.9

Table 2 Descriptive statistics for main variables in wage analysis

	All			Males		Females	
	Obs	Mean	Std. Err.	Obs	Mean	Obs	Mean
log wage	33,425	2.084	0.644	20,513	2.139	12,912	1.998
married (1 if married)	101,306	0.739	0.439	51,413	0.724	49,893	0.754
age	101,306	36.330	11.612	51,413	36.524	49,893	36.130
work experience (in years)	96,093	11.397	9.605	50,117	10.938	45,976	11.898
rural (1 if in rural area)	101,306	0.749	0.433	51,413	0.755	49,893	0.744
Red River Delta	101,306	0.186	0.389	51,413	0.184	49,893	0.189
North East	101,306	0.148	0.356	51,413	0.149	49,893	0.148
North West	101,306	0.056	0.230	51,413	0.055	49,893	0.057
North Central Coast	101,306	0.096	0.295	51,413	0.096	49,893	0.096
South Central Coast	101,306	0.087	0.281	51,413	0.087	49,893	0.087
Central Highland	101,306	0.068	0.252	51,413	0.070	49,893	0.067
Southeast	101,306	0.138	0.345	51,413	0.137	49,893	0.139
Mekong River Delta	101,306	0.220	0.414	51,413	0.223	49,893	0.217
No general education	101,306	0.187	0.390	51,413	0.162	49,893	0.212
primary (school grad.)	101,306	0.278	0.448	51,413	0.275	49,893	0.282
junior (high school grad.)	101,306	0.307	0.461	51,413	0.319	49,893	0.295
senior (high school grad.)	101,306	0.172	0.378	51,413	0.186	49,893	0.158
college (3–year)	101,306	0.015	0.120	51,413	0.012	49,893	0.018
university (4 or 5 year)	101,306	0.039	0.194	51,413	0.044	49,893	0.034
master (degree)	101,306	0.001	0.033	51,413	0.001	49,893	0.001
phd	101,306	0.000	0.015	51,413	0.000	49,893	0.000
other degree	101,306	0.001	0.024	51,413	0.001	49,893	0.000
no vocational degree	101,306	0.901	0.298	51,413	0.882	49,893	0.922
vocation 6 (months)	101,306	0.038	0.191	51,413	0.050	49,893	0.025
vocation 1 (year)	101,306	0.020	0.142	51,413	0.029	49,893	0.012
vocation 2 (year)	101,306	0.037	0.188	51,413	0.035	49,893	0.039
vocation 3 (year)	101,306	0.003	0.058	51,413	0.004	49,893	0.003
illiterate	101,306	0.064	0.245	51,413	0.049	49,893	0.080
family firms	96,093	0.005	0.070	50,117	0.005	45,976	0.005
family business	96,093	0.647	0.478	50,117	0.586	45,976	0.714
household business	96,093	0.157	0.364	50,117	0.209	45,976	0.102
state–owned firms	96,093	0.108	0.310	50,117	0.115	45,976	0.099
collectives	96,093	0.006	0.080	50,117	0.007	45,976	0.005
private firms	96,093	0.055	0.229	50,117	0.063	45,976	0.047
FDI firms	96,093	0.021	0.143	50,117	0.015	45,976	0.027

Table 3 Descriptive statistics for Panel 1

	All			Males		Females	
Variable	Obs	Mean	Std. Err.	Obs	Mean	Obs	Mean
enrollment	53,160	0.666	0.472	27,744	0.643	25,416	0.691
education (in years)	53,160	7.273	3.473	27,744	7.235	25,416	7.314
age	53,160	14.557	4.160	27,744	14.673	25,416	14.430
log household income	53,160	5.711	0.761	27,744	5.742	25,416	5.676
Head's self–employed	53,160	0.725	0.446	27,744	0.728	25,416	0.722
Head's in state–owned firms	53,160	0.056	0.230	27,744	0.055	25,416	0.057
Head's in collectives	53,160	0.006	0.076	27,744	0.006	25,416	0.006
Head's in private firms	53,160	0.028	0.165	27,744	0.028	25,416	0.028
Head's in FDI firms	53,160	0.004	0.064	27,744	0.004	25,416	0.005
Head spouse's self–employed	53,160	0.800	0.400	27,744	0.801	25,416	0.799
Head spouse's in state–owned firms	53,160	0.041	0.198	27,744	0.040	25,416	0.042
Head spouse's in collectives	53,160	0.002	0.045	27,744	0.002	25,416	0.002
Head spouse's in private firms	53,160	0.015	0.120	27,744	0.015	25,416	0.014
Head spouse's in FDI firms	53,160	0.006	0.078	27,744	0.006	25,416	0.006
Head's education	47,662	7.518	3.899	24,882	7.490	22,780	7.549
Head spouse education	51,833	6.491	3.996	27,061	6.461	24,772	6.523
rural	53,160	0.789	0.408	27,744	0.790	25,416	0.787
only son of the Head	53,160	0.202	0.386	27,744	0.387	25,416	0.000
grand child of the Head	53,160	0.081	0.272	27,744	0.078	25,416	0.084
age 7–11	53,160	0.261	0.439	27,744	0.258	25,416	0.264
age 12–15	53,160	0.300	0.458	27,744	0.291	25,416	0.310
age 16–18	53,160	0.246	0.430	27,744	0.242	25,416	0.249
age 19–22	53,160	0.194	0.395	27,744	0.209	25,416	0.177
Red River Delta	53,160	0.179	0.384	27,744	0.177	25,416	0.181
North East	53,160	0.146	0.353	27,744	0.145	25,416	0.148
North West	53,160	0.058	0.234	27,744	0.058	25,416	0.059
North Central Coast	53,160	0.118	0.322	27,744	0.118	25,416	0.118
South Central Coast	53,160	0.095	0.293	27,744	0.095	25,416	0.095
Central Highland	53,160	0.091	0.287	27,744	0.092	25,416	0.090
Southeast	53,160	0.129	0.335	27,744	0.129	25,416	0.128
Mekong River Delta	53,160	0.184	0.388	27,744	0.187	25,416	0.181

Table 4 Descriptive Statistics for Panel 2

	All			Males		Females	
Variable	Obs	Mean	Std. Err.	Obs	Mean	Obs	Mean
enrollment	37,646	0.940	0.237	18,978	0.940	18,668	0.941
education (in years)	37,646	7.106	3.395	18,978	7.004	18,668	7.210
age	37,646	12.985	3.536	18,978	12.920	18,668	13.050
log household income	37,646	5.758	0.759	18,978	5.786	18,668	5.729
Head's self–employed	37,646	0.715	0.452	18,978	0.717	18,668	0.712
Head's in state–owned firms	37,646	0.069	0.253	18,978	0.070	18,668	0.068
Head's in collectives	37,646	0.007	0.081	18,978	0.007	18,668	0.007
Head's in private firms	37,646	0.033	0.178	18,978	0.033	18,668	0.033
Head's in FDI firms	37,646	0.005	0.069	18,978	0.004	18,668	0.005
Head spouse's self–employed	37,646	0.789	0.408	18,978	0.791	18,668	0.787
Head spouse's in state–owned firms	37,646	0.053	0.224	18,978	0.053	18,668	0.053
Head spouse's in collectives	37,646	0.002	0.050	18,978	0.002	18,668	0.003
Head spouse's in private firms	37,646	0.017	0.128	18,978	0.018	18,668	0.016
Head spouse's in FDI firms	37,646	0.007	0.085	18,978	0.007	18,668	0.007
Head's education	34,077	8.165	3.803	17,250	8.177	16,827	8.152
Head spouse education	36,717	7.161	3.937	18,511	7.159	18,206	7.163
rural	37,646	0.770	0.421	18,978	0.772	18,668	0.768
only son of the Head	37,646	0.195	0.378	18,978	0.387	18,668	0.000
grand child of the Head	37,646	0.099	0.298	18,978	0.096	18,668	0.101
age 7–11	37,646	0.358	0.479	18,978	0.367	18,668	0.349
age 12–15	37,646	0.372	0.483	18,978	0.370	18,668	0.373
age 16–18	37,646	0.218	0.413	18,978	0.212	18,668	0.224
age 19–22	37,646	0.052	0.223	18,978	0.051	18,668	0.053
Red River Delta	37,646	0.201	0.401	18,978	0.202	18,668	0.200
North East	37,646	0.146	0.353	18,978	0.144	18,668	0.147
North West	37,646	0.050	0.218	18,978	0.053	18,668	0.046
North Central Coast	37,646	0.132	0.339	18,978	0.133	18,668	0.132
South Central Coast	37,646	0.103	0.304	18,978	0.103	18,668	0.103
Central Highland	37,646	0.088	0.283	18,978	0.087	18,668	0.089
Southeast	37,646	0.125	0.331	18,978	0.124	18,668	0.127
Mekong River Delta	37,646	0.154	0.361	18,978	0.153	18,668	0.155
ratio of tuition fee	37,521	0.115	0.160	18,910	0.117	18,611	0.113
education subsidies	37,646	0.098	0.298	18,978	0.097	18,668	0.100
scholarship	37,646	0.140	0.347	18,978	0.116	18,668	0.165
tuition fee reduction	37,646	0.419	0.493	18,978	0.420	18,668	0.418

Table 5 Wage function estimates

	(1)		(2)		(3)	
VARIABLES	log wage	Robust Std. Err.	select	Robust Std. Err.	Inverse mills ratio	Robust Std. Err.
sex	–0.184***	0.0191	–0.438***	0.0091		
age	0.0345***	0.0020	–0.0180***	0.0005		
age^2	–0.000428***	0.0000				
work experience	0.0235***	0.0013				
work experience^2	–0.000246***	0.0000				
rural	–0.0784***	0.0094	–0.313***	0.0108		
North East	0.0191	0.0166	–0.607***			
North West	0.113***	0.0296	–1.013***	0.0265		
North Central Coast	–0.0540***	0.0146	–0.428***	0.0181		
South Central Coast	0.0369***	0.0099	–0.0255	0.0180		
Central Highland	0.161***	0.0198	–0.625***	0.0213		
Southeast	0.227***	0.0091	0.103***	0.0159		
Mekong River Delta	0.0710***	0.0093	–0.133***	0.0146		
primary (school grad.)	0.0743***	0.0130	–0.0763***	0.0157		
junior (high school grad.)	0.0922***	0.0131	–0.0185	0.0161		
senior (high school grad.)	0.140***	0.0147	0.125***	0.0185		
college (3–year)	0.310***	0.0400	1.631***	0.0417		
university (4 or 5 year)	0.494***	0.0369	1.749***	0.0309		
master (degree)	0.779***	0.0709	1.996***	0.1950		
phd	0.573***	0.1220	7.304	–		
other degree	0.330***	0.0885	1.174***	0.1910		
sex*primary	–0.0255	0.0191				
sex*junior	0.0451**	0.0188				
sex*senior	0.0619***	0.0190				
sex*college	0.0949***	0.0326				
sex*university	0.0600**	0.0233				
sex*master	0.0695	0.0970				
sex*phd	0.416	0.2660				
sex*other degree	0.252	0.1740				
vocation 6 (months)	0.0194	0.0194	0.682***	0.0220		
vocation 1 (year)	0.0588**	0.0232	0.801***	0.0305		
vocation 2 (year)	0.0956***	0.0278	1.173***	0.0249		
vocation 3 (year)	0.115***	0.0380	1.126***	0.0780		
illiterate	–0.0459***	0.0161	0.0389*	0.0234		
state–owned firms	0.189***	0.0114				
collectives	–0.0479**	0.0204				
private firms	0.153***	0.0090				
FDI firms	0.296***	0.0141				
married			–0.164***	0.0117		
lambda					–0.0803**	0.0361
Constant	1.528***	0.0487	0.785***	0.0272		
Observations	101,306		101,306		101,306	

Standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Coefficients included but not displayed are 87 dummies for 88 industries where the employers have main business, and 33 dummies for occupational types. GSO classifies industries and occupational types in the survey.

Table 6 Gender gap in probability to leave school in Panel 1

	(1)		(2)	
VARIABLES	Haz. Ratio	Robust Std. Err.	Haz. Ratio	Robust Std. Err.
sex	0.766***	0.0192	0.840***	0.0192
log household income	0.608***	0.0192	0.786***	0.0118
Head's self–employed			0.999	0.0361
Head's in state–owned firms			0.877**	0.0458
Head's in collectives			0.966	0.124
Head's in private firms			0.999	0.0691
Head's in FDI firms			1.625***	0.219
Head spouse's self–employed			0.900***	0.0335
Head spouse's in state–owned firms			0.904	0.0567
Head spouse's in collectives			0.973	0.185
Head spouse's in private firms			0.964	0.0845
Head spouse's in FDI firms			1.094	0.147
Head's education			0.913***	0.00321
Head spouses' education			0.921***	0.00331
rural	1.245***	0.0277	1.052**	0.0272
only son of the Head	0.936***	0.0209	0.989	0.0241
grand child of the Head	0.819***	0.04	0.482***	0.0457
age 12–15	2.009***	0.233	1.902***	0.236
age 16–18	3.438***	0.39	3.176***	0.385
age 19–22	5.196***	0.589	4.292***	0.519
sex*age 7–11	1.119	0.181	0.893	0.16
sex*age 12–15	1.009	0.0543	0.945	0.0552
sex*age 16–18	0.984	0.0342	0.941	0.0366
North East	1.284***	0.0312	1.133***	0.032
North West	1.559***	0.0566	1.191***	0.0479
North Central Coast	0.947*	0.0283	0.959	0.0331
South Central Coast	1.172***	0.0365	0.846***	0.0316
Central Highland	1.758***	0.0587	1.228***	0.0467
Southeast	1.894***	0.057	1.158***	0.0407
Mekong River Delta	2.482***	0.0621	1.367***	0.0422
Observations	51,833		45,191	

Standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Model in use is Cox proportional hazard.

Table 7 Gender gap in probability to leave school in Panel 2

VARIABLES	(1)		(2)	
	Haz. Ratio	Robust Std. Err.	Haz. Ratio	Robust Std. Err.
sex	0.656***	0.054	0.688***	0.0598
log household income	0.843***	0.0275	0.928**	0.0342
Head's self–employed			1.041	0.0925
Head's in state–owned firms			0.970	0.106
Head's in collectives			1.198	0.282
Head's in private firms			1.305*	0.191
Head's in FDI firms			1.704	0.594
Head spouse's self–employed			1.255**	0.129
Head spouse's in state–owned firms			1.037	0.14
Head spouse's in collectives			1.055	0.449
Head spouse's in private firms			1.304	0.272
Head spouse's in FDI firms			1.442	0.558
Head's education			0.975***	0.00926
Head spouses' education			0.977**	0.00905
rural	1.122**	0.0566	1.022	0.0562
only son of the Head	0.916	0.0534	0.910	0.0566
grand child of the Head	0.918	0.107	0.899	0.175
age 12–15	0.146***	0.0345	0.133***	0.0353
age 16–18	0.0294***	0.00763	0.0311***	0.00882
age 19–22	0.0169***	0.00451	0.0172***	0.005
sex*age 7–11	1.327	0.333	0.940	0.27
sex*age 12–15	1.172	0.185	1.131	0.193
sex*age 16–18	1.373***	0.128	1.269**	0.125
North East	1.272***	0.0702	1.315***	0.0785
North West	1.117	0.108	1.182*	0.119
North Central Coast	0.989	0.0641	0.973	0.0684
South Central Coast	0.936	0.0691	0.877	0.0711
Central Highland	1.034	0.0917	1.004	0.0943
Southeast	0.775***	0.0682	0.738***	0.0708
Mekong River Delta	1.158**	0.076	1.007	0.0762
ratio of tuition fee over total expenditure on education	0.311***	0.0393	0.299***	0.0408
education subsidies	1.123	0.097	1.028	0.103
scholarship	0.700***	0.0506	0.708***	0.0573
tuition fee reduction	0.934	0.0561	0.929	0.0602
Observations	37,414		32,939	

Standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Model in use is Cox proportional hazard.

Table 8 Marginal effect of log wage over sex and degree obtained

Marginal effects	Males		Females	
sex#diploma	exp(b)	P>|z|	exp(b)	P>|z|
sex#no degree	7.832	0.000	6.413	0.000
sex#primary	8.288	0.000	6.534	0.000
sex#junior	8.339	0.000	7.122	0.000
sex#senior	9.419	0.000	8.225	0.000
sex#college	13.398	0.000	13.327	0.000
sex#university	18.028	0.000	16.312	0.000
sex#master	25.242	0.000	22.421	0.000
sex#phd	23.027	0.000	27.533	0.000
sex#upper degree – sex#lower degree	exp(b)	P>|z|	exp(b)	P>|z|
no degree–primary	1.579	0.000	1.128	0.226
primary–junior	1.052	0.520	1.801	0.000
junior–senior	2.945	0.000	3.015	0.000
senior–college	53.466	0.000	164.315	0.000
senior–univesity	5,478.685	0.000	3,251.152	0.000
university–master	1,358.479	0.000	449.724	0.000
master–phd	0.109	0.463	166.035	0.452

Table 9 Marginal effect of hazard rate to withdraw from school over sex and group of age in Panel 1, Analysis 1.

Marginal effects	Males		Females	
sex#age group	exp(b)	Pr>|z|	exp(b)	Pr>|z|
sex#age 7−11	0.106	0.000	0.095	0.000
sex#age 12−15	0.207	0.000	0.168	0.000
sex#age 16−18	0.340	0.000	0.271	0.000
sex#age 19−22	0.546	0.000	0.430	0.000
sex#upper age – sex#lower age	exp(b)	Pr>|z|	exp(b)	Pr>|z|
age 7–11 to 12–15	1.106	0.000	1.076	0.000
age 12–15 to 16–18	1.142	0.000	1.108	0.000
age 16–18 to 19–22	1.229	0.000	1.173	0.000