Returns to Higher Education in Vietnam: Causality and Heterogeneity

LE Thi Thuy Linh[1]

Paris-Dauphine University

IRD-DIAL

April 2015

Abstract: Higher education is highly appreciated by the public as well as policy-makers in Vietnam. The current boom of colleges and universities, however, has raised doubts on their quality. Using Labor Force Survey 2007, this paper aims at examining the impact of higher education on individual earnings in the labor market with a special focus on impact heterogeneity across the population. Applying OLS, Propensity Score Matching and IV methods, the author finds relatively large returns that greatly vary across individuals. Interestingly, a unique Heterogeneous Treatment Effects analysis reveals evidence for Negative Selection Hypothesis: those who are less likely to obtain tertiary education tend to benefit more from it.

Key words: returns to education, higher education in Vietnam, heterogeneous treatment effects, Negative Selection Hypothesis.

[1] Email : linh.le@dial.prd.fr and lethuylinh.bl@gmail.com.

1. Introduction

Higher education in Vietnam is highly appreciated by the public as well as policy-makers. On the one hand, there exists a conventional belief that higher education could mark a lifetime change, and everyone would benefit from it. Rural parents sometimes migrate to the cities like Hanoi to support their children's study by doing informal jobs such as food street vendors. On the other hand, higher education is a priority of Vietnamese government. To realize their ambitious plan of having one top-200 university and 20,000 PhDs by 2020, Vietnam has spent nearly one percent of their GDP on higher education, compared to 0.15%, 0.53% and 0.32% for Myanmar, Indonesia and the Philippines, respectively (UNESCO, 2014).

However, the current boom of colleges and universities has raised doubts on the quality of higher education in Vietnam. According to Valley and Wilkinson (2008), Vietnam lacks even a singly university of recognized quality, and about half of Vietnamese university graduates are unable to find jobs in their area of specialization. In this context, the returns to higher education in Vietnam should be accurately assessed to guide policies.

This paper aims at examining the impact of higher education on individual earnings in the labor market with a special focus on impact heterogeneity across the population. The author finds relatively large returns that greatly vary across individuals. Interestingly, those who are less likely to obtain tertiary education tend to benefit more from it. The paper is structured as follows. Section 2 reviews literature on returns to education on global and national bases. Section 3 presents data and empirical strategies, followed by statistical and econometrical findings in section 4. The last section concludes.

2. Literature Review

Return to higher education has been investigated in a vast literature all over the world. In general, it is found positive and significant, varying across the population (Card, 1999; Psacharopoulos & Patrinos, 2004). The world average private return to higher education is recorded at 19%, with the highest return for Sub-Saharan Africa (27.8%) and the lowest for OECD countries (11.6%), according to a survey by Psacharopoulos and Patrinos (2004). Astonishingly, it is documented that the estimation method makes little difference on the returns to education. Card (1999), after surveying a vast literature, concludes that the average return to education is not much below the estimate obtained from a standard human capital earnings function fit by OLS. Indeed, the instrumental-variable (IV) estimates are often 20%-40% higher than the corresponding OLS estimates (ibid).

Furthermore, numerous studies have found evidence for heterogeneity in returns to education, such as Card and Krueger (1992), Blundell et al. (2004), Heckman and Li (2004). In fact, it is possible that the impact of education varies according to some individual characteristics or with treatment probability. There are two opposite views on this. On the one hand, rational-behavioral models consider that college decision is made based on a cost-benefit consideration, suggesting that individuals who are the most likely to take higher education would also benefit the most from it (Mincer, 1974; Carneiro et al., 2001; Heckman & Li, 2004). On the other hand, there may be numerous non-economic factors that influence higher education decision as being acknowledged in sociological literature. College attainment is driven not only by rational choice but also by cultural and social norms and circumstances (Coleman 1988, as cited in Brand & Xie 2010). In fact, Brand and Xie (2010) uncovers a surprising *negative selection* phenomenon, i.e., those who are least likely to obtain a college education benefit the most from it. Consequently, programs that raise the educational level of children from poor family backgrounds tend to have higher marginal returns (Kane & Rouse, 1995).

In Vietnam, returns to education have been estimated in a number of studies. Despite the difference in magnitude, the impact of education appears to be always positive (Doan, 2011; Doan & Gibson, 2009; Liu, 2006; Arcand et al., 2004; Moock et al., 2003; Glewwe & Patrinos, 1999; Stroup & Hargrove, 1969). While Liu (2006) reports a downward trend from 1992 to 1998 in the rates of return for men, Doan and Gibson (2009) found a dramatic rise in returns to schooling over the 1998-2004 period. In 2008, university education is expected to raise individual wage rate by 68%, *ceteris paribus* (Doan, 2011), which is much larger than the average level of Asia (18.2%, Psacharopoulos & Patrinos, 2004). This could be attributed to recent labor market reforms toward openness and international integration of Vietnam's economy.

Yet, Moock et al. (2003) points out a remarkable heterogeneity in the returns to education with respect to sex, level of education and institutional sector. Females experienced a much higher rate of return than males. Workers in the public sector have higher private rates of returns to education than do private sector counterparts. The authors suggest that there are factors other than education, for example Communist Party membership, that distort public sector pay.

These aforementioned research work encounter several important limits. The ***first*** shortcoming lies in non-representative samples used in the cases of Stroup and Hagrove (1969) and Moock et al. (2003). Stroup and Hargrove (1969) covers rural South Vietnam only. Moock et al. (2003) only investigates the wage workers, who accounted for only 20% labor force at that time. It is emphasized that *"the estimates of the returns to schooling could be much higher than estimated here if the self-employed were included,*" and *"it could be that in the less tightly regulated informal sector, productivity has a higher payoff."*

Second, a number of studies only cover spurious correlation between education and earnings (Liu, 2006; Moock et al., 2003; Glewwe & Patrinos, 1999; Stroup & Hargrove, 1969). In Vietnam, higher education decision is influenced by liquidity constraints (Glewwe & Jacoby, 2004; Glewwe & Patrinos, 1999) and entrance examinations. Consequently, factors such as family resources, individual ability and motivation, if being ignored in returns-to-education regressions, would cause omitted variable bias.

Third, although Moock et al. (2003) has found a heterogeneity in returns to education; unfortunately, all of previous studies have blocked themselves in a Homogeneous Treatment Effect framework. Given that heterogeneity, even standard IV method is unable to produce an unbiased estimate of returns to schooling. After conducting a series of tests, Arcand et al. (2004) concludes that few IVs met the two validity conditions. Among their list, only parental education and the matrix of instruments proposed by Hausman and Taylor (1981) are satisfactory, with the latter being able to solve the problem of LATE.

This paper overcomes the aforementioned limits as follows. *First*, the data used is representative for the whole population of Vietnam. This 2007 Labor Force Survey is indeed the first phase of an *1-2-3 scheme*, which allows to identify informal employment in accordance with international definition. This enables analyses focusing on the informal sector. *Second*, the causal impact of higher education is attempted to identify by using various econometrical techniques. Particularly, month of birth is used as an instrument variable. *Last but not least*, this research work investigates heterogeneous treatment effects, through separate regressions on different sub-populations, by allowing interaction terms between higher education and individual characteristics, and notably by applying a unique "Heterogeneous Treatment Effect – HTE algorithm" developed by Jann (2010) and Xie et al. (2012).

3. Data and methodology

To the best of my knowledge, this is the first paper that examines the returns to schooling based on Vietnam Labor Force Survey (LFS 2007), a rich and nationally representative dataset. Most of previous literature have exploited the Vietnam Living Standards Survey (Doan, 2011; Doan & Gibson, 2009; Liu, 2006; Arcand et al., 2004; Moock et al., 2003; Glewwe & Patrinos, 1999). LFS 2007 covers up to 173,000 households (IMF),[2] compared to only 4,800 households in VLSS 1992-93 and 45,945 households in VHLSS 2006. Large sample size, on the one hand, offers more precise estimates. On the other hand, that facilitates separate analyses on specific sub-populations, such as informal workers. In fact, LFS 2007 contains necessary information for the identification of internationally-defined informal sector and informal employment.

One may argue that one advantage of VLSS over LFS is that VLSS supports panel data analyses, while LFS does not. Veritably in this particular setting, higher education as a treatment status offers little variation from year to year. Neither LFS nor VHLSS provide exact number of years of higher education. Moreover, given limited time span, those who change their treatment status are either newly graduates or officials taking in-service training (« tai chuc ») of usually low quality. The returns to higher education for these groups might be specific. Furthermore, measurement error in schooling tends to be higher in fixed-effect model than in cross-sectional estimator (Ashenfelter & Zimmerman, 1997; Belzil, 2007 as cited in Doan, 2011). These are the reasons why cross-sectional analysis is chosen over panel one in this research work.

Empirical Strategies

In a non-experimental context, the main challenge to causal effect identification is that one cannot observe the earnings of the same individual in both states: with and without going to college. Simply comparing university graduates with the rest leads to a certainly biased estimate because these two groups are not the same. Therefore, an appropriate identification strategy is needed to estimate the causal link. Blundell et al. (2004) stress that "no given non-experimental estimator is uniformly superior to all others". Indeed, both OLS and matching estimates depend on the same crucial assumption of selection on observables, and "both are thus as good as the

[2] http://dsbb.imf.org/pages/gdds/DataCategoryReport.aspx?ctycode=VNM&catcode=LEMP0

control variables". Likewise, the credibility of IV and control function approaches all rely on an exclusion restriction (ibid).

A naive bivariate Ordinary Least Square (OLS) regression of individual income on her schooling is written as follows:

$$y_i = \beta_0 + \beta_1 S_i + \varepsilon_i$$

Where y_i stands for natural logarithm of individual earnings from wage, bonus and subsidies; $S_i$is higher education dummy.[3] S_i equals 1 if one has finished college education and above, and 0 if she has only obtained high school diploma.

This econometric model would potentially suffer from following sources of bias:

(1) ***Omitted variable bias (Pre-treatment heterogeneity bias)***. Under the restriction that β_1 is homogeneous across individuals, the main threat to causal inference of an OLS regression is a non-zero correlation between ε_i and S_i. There are some factors that drive both educational decision of individuals and their earnings in the labor market such as gender, ethnicity, region, ability and parents' encouragement. For example, more capable people, who would earn more in the labor market *ceteris paribus*, are also more likely to obtain more education. The omission of ability thus possibly leads to an overestimation of the returns to schooling.

(2) ***Measure errors***. While measurement error in earnings only worsens precision of the model, measurement error in schooling causes attenuation bias. In fact, prior research has generally reported the reliability of self-declared schooling at only 90% (Card, 1999). Fortunately, such a downward bias could be balanced with aforementioned ability upward bias.

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

If the impact of education varies according to some individual characteristics or with treatment probability, conventional estimates of returns to schooling might be biased. Imbens and Angrist (1994) state that an IV estimate only recovers a local treatment effect, i.e., the effect on those whose educational decision is affected by the instrument ("compliers"). Moreover, Heckman and Li (2004) show that OLS gives a downward biased estimate and IV produces an upward biased estimate of ATE.[4] The disparity between IV and OLS estimates depends on the extent that instruments affect schooling decision at different education levels due to heterogeneous returns to schooling (Card, 1999). For example, IV estimates based on instruments which

[3] "There is considerable evidence in many contexts that returns to schooling are nonlinear in years of schooling, so conventional log wage on years of schooling regression coefficients generate rates of return that are badly biased estimates of the return to college education" (Heckman et al., 2003 as cited in Heckman and Li, 2004).

[4] They adopt a Marginal Treatment Effect (MTE) concept, which is introduced by Bjorklund and Moffitt (1987). MTE is "the average return to schooling for persons indifferent to going on to schooling at different levels of unobservable factors that determine schooling choices."

influence schooling decisions of children from relatively disadvantaged family background (e.g. lower parental education, income, assets) tend to be higher than the OLS estimates.

There exist two opposite findings on the trend of variation in returns to education across individuals. As mentioned earlier, both Positive Selection Hypothesis and Negative Selection Hypothesis are supported by empirical literature (Carneiro et al., 2001; Heckman & Li, 2004; Brand & Xie, 2010).

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

In order to single out the causal impact of education on earnings, it is a must to correct for these biases.

First, some covariates $X_i's$ would be added into the model to reduce omitted variable bias. $X_i's$ consist of individual observables, which must be exogenous in the sense that they are unaffected by schoolings. Ideal candidates are pre-treatment characteristics and time-invariant factors such as gender, ethnicity and region.[5] In contrast, sector indicators are "bad" controls as one benefit of education is that it facilitates sectorial choice, thus including them in the earnings equation might underestimate returns to schooling (Heckman & Li, 2004). Control variables could also be used to match "treated" and "untreated" individuals in propensity score matching. Nevertheless, confounding factors like ability cannot be *perfectly* controlled for due to data unavailability. At best, individual ability and family background is proxied by parental income or parental education. Fortunately, the potential upward bias caused by the unobservable may be partially offset by the attenuation measurement error bias (Blundell et al., 2004).

Furthermore, even under selection on the unobserved, the causality could be recovered thanks to the exploitation of some 'exogenous' variation in education by using an excluded instrument. A good instrument must be significantly correlated with individual schooling but uncorrelated with the residuals of regression equation. In other word, it affects individual earnings only through education.

Second, month of birth is employed as an excluded instrument in the analysis. On the one hand, month of birth is naturally exogenous. On the other hand, it may affect individual schooling via several channels (Crawford et al., 2007; Angrist & Krueger, 1991). In Vietnam, school years start in every September and children start going to school the year when they turn six. Children

[5] *Region* (rural-urban area) is included in the model, keeping in mind that migration potentially "endogenises" this variable. Differently speaking, one's current living location is not necessarily their hometown, and probably affected by their education level.

born toward the end of a year are therefore less mature and less prepared than their peers who are born toward the beginning of a year. Consequently, the latter might better perform and finally achieve higher qualifications. However, there may be a tendency that parents of the former keep their child at home till the next school year (i.e., when the child turns seven) in order to overcome the aforementioned disadvantage. Meanwhile, children born at the beginning of a year are more likely to join the previous cohort, i.e., they start going to school at the age of five.[6]

Third, to account for potential heterogeneous returns to schooling, separate regressions for different sub-sample which may have different returns are performed. In addition, interactions between higher education dummy and several explanatory variables would reveal which individual characteristics are more associated with higher returns to higher education. Remarkably, a propensity score-based heterogeneous treatment effect analysis (so called "HTE" algorithm) is conducted to figure out the pattern of treatment effects across propensity score strata (Jann, 2010; Xie et al., 2012). Nonetheless, one should keep in mind that this analysis is based on the *ignorability assumption, i.e. "selection on observables"* (Brand &Xie, 2010).

To be detailed, the *« Stratification-Multilevel method»* (Xie et al. 2012) includes four steps as follows:

i) Estimation of propensity score by using a *Probit* model.

ii) Construction of balanced propensity score strata where there are no significant differences in the average values of covariates and the propensity score between the treatment and control groups.

iii) Estimation of strata-specific average treatment effects.

iv) Estimation of the trend of treatment effects across propensity score strata.

4. Findings

4.1. Descriptive Statistics

Table 1 provides a comparison of those who obtain higher education and those who do not in terms of demographic as well as current employment characteristics. In general, the former group have more privileged background and enjoy better working conditions than the latter.

Regarding demographic characteristics, those with tertiary education are less likely to belong to ethnic minority and rural areas than those without. Their family background is also more favorable, with much wealthier parents and a substantially higher proportion of parents having completed higher education compared to the non-treated group (66% and 40% versus 35% and 27% for father and mother, respectively). Meanwhile, there is inconsiderable gender difference between the two groups, indicating apparently equal opportunities of benefiting from higher education between males and females. Notably, the treated group are slightly older than the

[6] What is observed with Vietnam LFS 2007 is such a so called *reverse maturity effect*.

non-treated counterpart (40 versus 34 years old). Thus, there might be a possibility that some have not obtained university degree only because they are not old enough.[7] However, while taking only those with available wage information into account, the age profiles of two groups become comparable (37 years old for treated group and 36 years old for control group).

Unsurprisingly, better educated group also acquire more positive employment characteristics. First, their average earnings level is tremendously larger than that of high school graduates (11.47 compared to 7.25 per hour, in thousand Vietnam dongs). A majority of college graduates work for public sector (68%), and this proportion is almost three times as much as that of the non-treated group (24%). In Vietnam, public sector employment is highly appreciated for their stability and social status. Logically, while a large part of high school leavers are working for the informal sector, only 15% of university graduates engage in this vulnerable employment category. Almost all of the high qualified are wage workers (88%), whilst more than half of the less qualified work under other employment status such as self-employed and family worker.

Table 1. Descriptive Statistics of Precollege Covariates and Employment Characteristics

Variable	Non-treated	Treated
Ethnic minority	6.78%	3.72%
Female	47.38%	47.20%
Rural area	54.70%	30.66%
Age	34.37	39.63
Father completing higher education	35.43%	65.60%
Mother completing higher education	27.48%	39.58%
Father's income (log)	1.06	1.92
Mother's income (log)	0.24	0.88
Hourly income	7.25	11.47
Informal employment	62.43%	14.55%
Public sector	23.95%	68.46%
Wage worker	47.94%	87.90%

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

4.2. Estimates of Returns to Higher Education

4.2.1. OLS, Matching and IV

Table 2 presents OLS estimates of the impact of higher education on individual earnings in Vietnam. In general, the return to higher education in Vietnam is positive and statistically significant. On average, university graduates earn 60% more than high-school leavers, *ceteris paribus*. When age is additionally controlled for, the return to higher education slightly decreases to be 58%. This reveals that age might cause an upward bias if it stays in the error term. Interestingly, the estimate noticeably changes when a proxy for ability and family

[7] Indeed, if the sample is restricted within more-than-24-year-old people, the average age of two groups are almost equal. By the age of 24, normally in Vietnam one should have finished her university education. Social sciences bachelor programs often last four years, engineering programs five years and medical courses the longest – six years.

background is added into the model. While adding father's income into the control list, the estimated return to higher education lightly raises to 65%, which is close to Doan (2011)'s estimate at 68%. In contrast, while ability and family background are captured by mother's education, the estimate significantly drops to 38%. This supports the conventional belief of upward ability bias in earnings equation. In all specifications, the estimated returns are excessively higher than those of Moock et al. (2002) at around 10%. This could be attributed to the fact that the sample used in Moock et al. (2002) is far from representative for Vietnamese population: it excludes up to 80% of the labor force at that time. Moreover, profound labor market reforms since 1990s might have caused remuneration increasingly responsive to workers' educational level. Liu (2006) indeed has proved that wage structure has changed in such a way that favors more educated workers during Vietnam's transition.

Table 2. Returns to Higher Education in Vietnam - OLS Estimates

	I	II	III	IV	V	VI
	0.601***	0.583***	0.494***	0.570***	0.647***	0.383***
	(0.007)	(0.007)	(0.01)	(0.007)	(0.176)	(0.119)
Minor ethnicity, female, rural (X)	x	x	x	x	x	x
Age (logarithm)		x	x	x	x	x
Interaction of (X) with HE dummy			x			
Province dummies				x		
Father's income					x	
Mother's education						x
Number of observation	100028	100028	100028	100028	71	114

*Note: *** 1% ; **5% ; *10% level of significance.*

Although controlling for personal ability and family background could help reduce potential bias, the inclusion of this information into the model enormously contracts the sample size. Actually, information on parental income and education is available only for household heads who live together with their parents in a same household, and whose parents are willing and capable to answer the questionnaire. Consequently, this subsample is hardly representative for the whole sample. To examine the representativeness of subsamples with available information on parental income and education, principal demographic and employment characteristics of these subsamples are put in comparison with those of the whole sample in Table 3.

Unsurprisingly, these subsamples differ from the whole sample in various aspects. There is a much less percentage of females in these subsamples (10% and 27%), which could be attributed to the fact that in a Confucianism-affected country like Vietnam, household heads who live with their old parents are usually men. Moreover, these two subsamples are also dissimilar from each other, explaining the disparity in return-to-education estimates obtained from models (V) and (VI). Subsample (1) has a much higher proportion of ethnic minority, rural area residents and informal workers, but a notably lower percentage of public sector workers, wage workers and smaller earnings level than subsample (2). Therefore, a larger impact of higher education on earnings recorded for subsample (1) compared to subsample (2) to some extent reveals that the disadvantaged labour market players might benefit more from college than the advantaged.

Table 3. Representativeness of samples with available information on parental income and education

Variable	**Whole sample**	**Father's income (1)**	**Mother's education (2)**
Ethnic minority	14.12%	33.62%	10.29%
Female	51.24%	10.44%	27.35%
Rural area	72.83%	84.63%	31.66%
Age	31.31	33.66	36.4
Hourly income	5.66	4.80	10.70
Informal employment	82.87%	82.84%	43.53%
Public sector	10.53%	10.33%	31.66%
Wage worker	30.03%	36.52%	61.72%

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

To further investigate this phenomenon, higher education dummy is allowed to interact with pre-treatment demographic factors, namely *minor ethnicity*, *female* and *rural*. As expected, all three interaction terms are positive and statistically significant, supporting the aforementioned hypothesis. In addition, separate regressions are run for each subsample as in Table 4. The regression results indicate that returns to higher education of the less advantageous groups such as females, ethnic minority, rural workers, are larger than the population average. For example, on average, finishing college is expected to raise hourly earnings of females by 63%, *ceteris paribus*, versus 57% for males. The earnings effect of university education is exceptionally important for minor ethnic workers and rural residents, at 77% and 86%, respectively.

Notably, public sector workers are less rewarded from higher education than the population average (35%), which is in opposition to Moock et al. (2003)'s finding. This might be justified by wage rigidity in this sector; however, its interpretation should be made with caution. In Vietnam, a prerequisite criterion to occupy senior positions in many public institutions is tertiary qualification. Although public pay is far from generous, university graduates could take advantage of their network and information to earn extra money which is hardly captured in the survey.

The return to higher education for informal workers stays at fairly low level, ranging from 23% to 33% according to estimates. This again disproves Moock et al. (2003)'s prediction, that productivity may have a higher payoff in a less tightly regulated informal sector. In fact, unlike public sector, informal sector in Vietnam is dominated by the self-employed (Cling et al., 2010), for those there is no payment escalator depending on educational level. Moreover, given precarious operating conditions and generally low earnings and poor labor conditions in this sector (ibid), it is plausible that university degree makes little difference in individual earnings. Nevertheless, it is totally possible that informal workers are more likely to hide their real income which is very hard to be verified.

IV estimates of returns to higher education in Vietnam are also reported in Table 4. Astonishingly, when month of birth is used as an excluded instrument, the estimated return turns out to be negative and significant at 10% level. The magnitude is unreasonably large: On average, undertaking college education is expected to reduce individual earnings up to 300%. This raises doubts on the validity of month of birth as an instrumental variable. First, although month of birth is naturally exogenous in theory, the declared month of birth may be not. In Vietnam, where regulations are weakly enforced, parents could misreport their child's birthday, and one could even modify their official birthday for different purposes. In fact, while the correlation between month of birth and individual education is found positive in Europe (England - Crawford et al., 2007; France – Grenet, 2010; Sweden - Fredriksson & Öckert, 2005), what is observed in Vietnam's LFS 2007 is a negative relationship. Second, month of birth explains a minimal fraction of individual probability of acquiring tertiary education (i.e., the first-stage R-squared is less than one percent), making it a potentially weak instrument. Indeed, as the sample is narrowed, the first-stage estimate becomes no longer significant. For these reasons, the obtained IV estimates might not be trustworthy.

4.2.2. Heterogeneous Treatment Effect

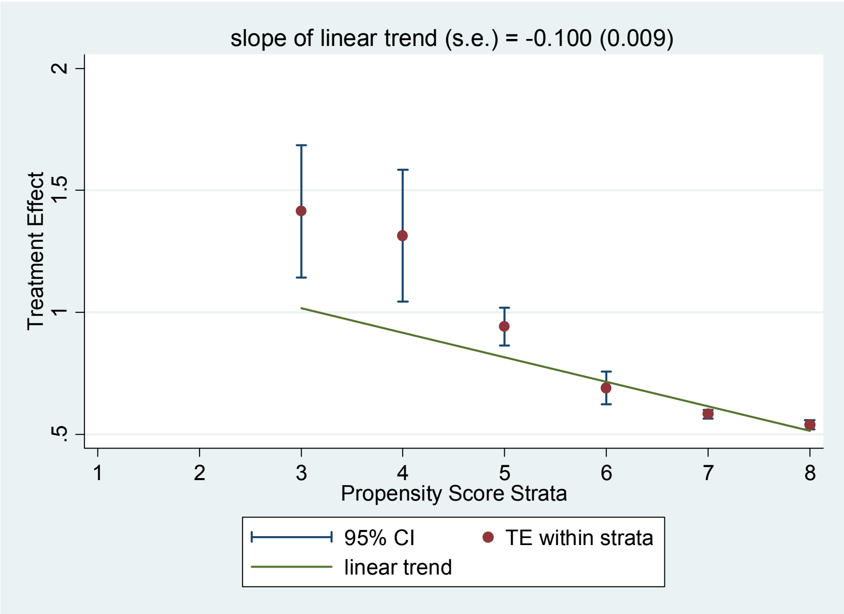
Previous estimation has discovered preliminary findings on heterogeneity in returns to higher education among different subpopulations in Vietnam. On the one hand, whether in-born disadvantaged individuals could better profit from college needs to be verified through a sophisticated algorithm. On the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

First, in order to learn the demographic profile of college graduates, a Probit regression is used to explain higher education dummy as a function of individual characteristics. *Ceteris paribus*, belonging to ethnic minority and rural areas instead of being *Kinh* and urban resident decreases the probability of going to university by 8% points and 42% points, respectively.[8] Meanwhile, gender has no significant impact on the possibility of obtaining higher education of individuals. These findings correspond to descriptive statistics as shown in Table 1. Given that there should be no intelligent difference among ethnic groups and different regions, this reflects excessive constraints that prevent the under-privileged from receiving college education. Besides liquidity constraint which is the most acknowledged in the literature as well as in policies, there could be other infrastructural and cultural hurdles. Minor ethnic groups, for examples, often live in remote mountainous areas with poor traffic conditions, sparse schools, little information, early marriage custom, whose people lack a recognition of education's benefit.

The next steps of SMM yield a suggestive pattern in treatment effects across propensity score strata. Figure 1 shows that those who are less likely to take higher education actually benefit more from it. To be specific, ethnic minority and rural residents are rewarded the most in terms of earnings from higher education, while they have lowest probability of going to college. Even when personal ability and family background are controlled for, a negative sloped trend still emerges (Figure 2). This result supports the Negative Selection Hypothesis proposed in sociological literature and confirmed by Brand and Xie (2012).

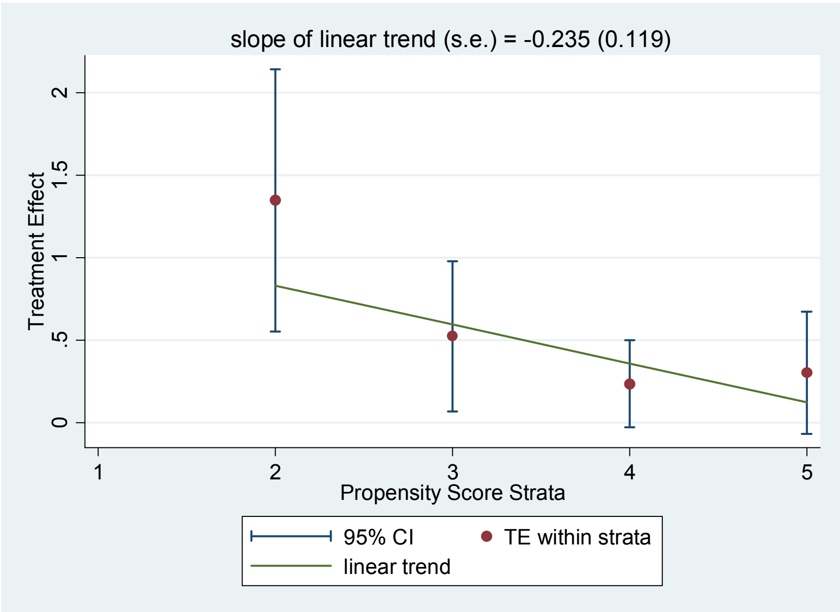
[8] Ethnic minority and rural areas are correlated to each other, with a covariance of 0.24. When probit regression is run on each variable separately, the coefficient magnitude turns out to be 16% and 51%, respectively.

Figure 1. Heterogeneous Treatment Effects of Higher Education in Vietnam



Note: Author's computation using STATA command "*hte sm*". Variables included to construct propensity score strata: *meth female rural*

Figure 2. Heterogeneous Treatment Effects of Higher Education in Vietnam, controlling for personal ability and family background



Note: Author's computation using STATA command "*hte sm*". Variables included to construct propensity score strata: *meth female rural motherhedu.*

Brand and Xie (2010) argue that the relatively large return to higher education of low-propensity university graduates could be explained by the fact that their social position is marked with substantial disadvantage. Without a college degree, these individuals have limited human and social capital that disables their labor market prospects. In the contrary, in the absence of university qualification, individuals from more advantaged social backgrounds could still benefit from their superior resources.

5. Conclusion

Using a large nationally representative dataset, this paper has made various attempts to quantify the causality of higher education and individual earnings and its heterogeneity across Vietnamese population. In general, the returns to higher education in Vietnam appear to be large according to international and regional standards. The estimated returns substantially exceed those of the 1990s (Moock et al., 2003), revealing a positive effect of labor market reforms since then. Interestingly, supporting evidence for Negative Selection Hypothesis is found: those who are less likely to go to college tend to benefit more from it with respect to earnings. This evidence suggests policies that encourage the disadvantaged to pursue higher education such as: bonus grade for ethnic minority and rural habitants in university entrance examination, financial grants or loans for underprivileged students and so forth.

Nonetheless, several limits of this research work could be addressed in further analyses. First, higher education has been tackled as a homogeneous category, since LFS 2007 data neither allow to distinguish different majors nor capture heterogeneous university program quality (ranking). Actually, some majors seem more promising than others: newly-graduated engineers, for example, may earn more than early-career teachers; English and Mathematics teachers are probably better rewarded than History or Geography ones. Likewise, recruiters may welcome Foreign Trade University students more than other less known schools. It would be interesting to incorporate this kind of information in the earnings equation. Second, although the initial idea of including personal ability and family background proxies as control variables as well as of using month of birth as an instrument sounds promising, their application poses a number of practical issues. Hopefully, future data of more accuracy and larger response rate would bring more reliable estimates.

Table 4. Returns to Higher Education in Vietnam for Various Subsamples

	All	Males	Females	Ethnic minority	Rural residents	Wage-workers	Non-wage workers	Public sector	Informal sector
Basic specification (BS)	0.601***	0.573***	0.631***	0.770***	0.861***	0.439***	0.380***	0.351***	0.326***
	100028	51950	48078	7573	17568	66029	33923	47659	37492
BS + age	0.583***	0.544***	0.627***	0.694***	0.802***	0.402***	0.237***	0.343***	0.283***
	100028	51950	48078	7573	17568	66029	33923	47659	37492
BS + age + province	0.570***	0.527***	0.615***	0.713***	0.722***	0.403***	0.202***	0.354***	0.236***
	100028	51950	48078	7573	17568	66029	33923	47659	37492
Matching: BS	0.593***	0.566***	0.621***	0.743***	0.860***	0.436***	0.381***	0.348***	0.324***
	100028	51950	48078	7353	17568	66029	33923	47659	37492
IV without controls	-3.132*	-1.322	x	x	x	-1.408	x	x	-11.319
	99566	51678	47888	7543	17476	65775	33715	47481	37297
IV with controls	-3.173*	-1.402	x	x	x	x	x	x	-8.153
	99563	51676	47887	7543	17476	65773	33714	47479	37297

Note: *** 1% ; **5% ; *10% level of significance. Number of observations under the estimates. "x" signals that the first stage of IV regression is not statistically significant. Basic specification is OLS with *gender, ethnicity* and *rural* as explanatory variables

References

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

Angrist, J. and Krueger, A. (1998). Empirical strategies in labor economics. In: O. Ashenfelter and D. Card, ed., *Handbook of Labor Economics*. Alsevier Science B. V., pp.1277-1366.

Arcand, J-L., d'Hombres, B., Gyselinck, P. and Nguyen Tuan, L. (2004). Instrument Choice and the Returns to Education: New Evidence from Vietnam. Working Papers No 200422, CERDI.

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

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

Card, D. (1999). The causal effect of education on earnings. In: O. Ashenfelter and D. Card, ed., *Handbook of labor economics*. Elsevier Science B. V., vol. 3, pp.1801-1863.

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

Carneiro, P., J. Heckman and E. Vytlacil (2001). Understanding IV: Estimating the Return to Education when It Varies Among Individuals. Working Paper, University of Chicago.

Cling, J., Nguyen, T., Nguyen, H., Phan, T., Razafindrakoto, M. and Roubaud, F. (2010). *The Informal Sector in Vietnam: a Focus on Hanoi and Ho Chi Minh City*. Hanoi: The Gioi Editions.

Crawford, C., Dearden, L. and Meghir, C. (2007). *When you are born matters*. London: CEE.

Doan, T. (2011). Labour Market Returns to Higher Education in Vietnam. *SSRN Journal*.

Doan, T. T., and Gibson, J. (2009). Do returns to schools go up during transition? The not so contrary case of Vietnam.

Glewwe, P. and Jacoby, H. (2004). Economic growth and the demand for education: is there a wealth effect?. *Journal of Development Economics*, 74(1), pp.33-51.

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

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

Heckman, J. and Li, X. (2004). Selection bias, comparative advantage and heterogeneous returns to education: evidence from China in 2000. *Pacific Economic Review*, 9(3), pp.155-171.

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

Jann, B. (2010). *Heterogeneous Treatment Effect Analysis*. Available at: http://www.stata.com/meeting/germany10/germany10_jann.pdf [Accessed 21 Apr. 2015].

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

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

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

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

Psacharopoulos, G. and Patrinos *, H. (2004). Returns to investment in education: a further update. *Education Economics*, 12(2), pp.111-134.

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

UNESCO (2014). *Higher Education in Asia: Expanding Out, Expanding Up: the Rise of Graduate Education and University Research*. [online] UNESCO Institute for Statistics. Available at: http://www.uis.unesco.org/Library/Documents/higher-education-asia-graduate-university-research-2014-en.pdf [Accessed 21 Apr. 2015].

Valley, T. and Wilkinson, B. (2008). Vietnamese Higher education: Crisis and response. *Memorandum: Higher Education Task Force, Harvard Kennedy School*.

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