Economics
The Open-Access, Open-Assessment E-Journal

Discussion Paper
No. 2011-4 | March 14, 2011 | http://www.economics-ejournal.org/economics/discussionpapers/2011-4

Labour Market Returns to Higher Education in Vietnam

Tinh Doan
University of Waikato, Hamilton

Please cite the corresponding journal article
www.economics-ejournal.org/economics/journalarticles/2011-12

Abstract This paper employs the Ordinary Least Squares, Instrumental Variables and Treatment Effect models to a new dataset from the Vietnam Household Living Standards Survey (VHLSS) to estimate return to the four-year university education in 2008. Our estimates reveal that the return to university education is about 17% (annualized) and robust to the various estimators. The return to higher education has significantly increased since the economic reform in late 1980s.

Paper submitted to the special issue
Quasi Markets in Education

JEL C31, J31, O15
Keywords Economic transition; returns to higher education; IV model; Vietnam

Correspondence Tinh Doan**,** Department of Economics, Waikato Management School, The University of Waikato, Hamilton, New Zealand; e-mail: thanhtinhdoan@gmail.com

1. Introduction

The most challenging task of estimating education treatment effect or return to education is that one does not have sufficient information about studied subjects. Observationally identical individuals make different choices; we do not know why some people decide to take the four-year university education, while some do not so that difference in their earnings may be affected by observed, unobservable attributes and education participation (treatment). To estimate return to the four-year university education, one should measure how much people would have earned if they did not have the four-year university degree (Heckman & Li, 2004). One is unable to measure the later earnings (counterfactual earnings).

The ordinary least squares (OLS) does not account for the factors affecting the four-year university schooling decision, and especially investment in education in Vietnam is faced with liquidity constraints (Glewwe & Jacoby, 2004; Glewwe & Patrinos, 1999). Furthermore, the four-year university entry is not free from competition due to the government's limited number of students and due to facility and human capacity of education providers (universities); about three fourths of high school leavers are unable to go to university (about 1.2 million student complete high school education in 2009).[2]

Only 2% of the Vietnam population move into higher education in 2009, it is much lower than regional and international context,[3] and only 5% of the population had ever attended university and post-graduate education (GSO, 2010, p. 8). Given the fact that university candidates have to complete high schools and take entrance examinations as well as face liquidity constraints, factors such as individual ability, and family resources and motivation may play important roles in their pursuing university education. Therefore, entering the four-year university education is agents' selectivity/competition by both family and students.

Our results show that the return to the four-year university education in Vietnam in 2008 is, on annually average, 17% based on IV model and 17.8% based on the OLS and Treatment Effect models. Thus, the bias by OLS model is not too large to be concerned in the context of higher education in Vietnam. The return to university education has remarkably improved after more than twenty years of economic transition. Given the fact that the return to education in Vietnam was very low in early 1990s (Glewwe & Patrios, 1999; Moock, Patrios, & Venkataraman, 2003), labour

---

[2] See at http://www.business-in-asia.com/vietnam/education_system_in_vietnam.html; and http://www.gso.gov.vn/default_en.aspx?tabid=474&idmid=3&ItemID=10220
[3] See at http://www.business-in-asia.com/vietnam/education_system_in_vietnam

market in Vietnam has begun to function more effectively, so higher-qualified labourers have been rewarded more than in the past.

This paper is structured as follows. Section 2 reviews literature on estimating methods for return to education. Section 3 presents empirical models and data. Section 4 discusses estimation results. Concluding remarks are presented in Section 5.

2. Literature review

To estimate the returns to schooling, the Mincerian earnings equation is the first point to begin, the model is as follows:

$$\text{Log } y_i = \lambda S_i + \beta X_i + u_i \tag{1}$$

When using the ordinary least squares (OLS) to estimate λ, one assumes that S is uncorrelated with the unobserved disturbance u_i of equation (1), but this may not be true. Estimated λ may be biased since individual's ability and motivation affect both earnings and education (Ashenfelter, Harmon, & Oosterbeek, 1999). More-able and higher-motivated individuals will stay at school longer and also earn more. Therefore, there is a debate about endogeneity of schooling decision which is not independent of other factors affecting earnings such as unobserved individual ability and motivation (Griliches, 1977). The measured correlation between education and earnings may not be a truly causal effect relationship. A part of earnings would result from ability that also affects education. This draws researchers' attention to overcoming this problem by employing IV method, twin and sibling data, and fixed effect model (Angrist & Krueger, 1991; Staiger & Stock, 1997; Card, 1995; Ashenfelter & Rouse, 1998; Miller et al, 1995; Ashenfelter & Zimmerman, 1997; Butcher & Case, 1994; Hausman & Taylor, 1981).

The main point of attention is that education is not randomly assigned to individuals, and their choices are heavily reliant on many factors such as their ability, motivation and family background (Card, 1995). Card (1995, 1999, and 2001) suggests careful rethinking of factors influencing schooling decision. Education attainment may be endogenous, and hence earnings equation is postulated as follows:

$$\text{Log } y_i = \lambda S_i + \beta X_i + u_i \tag{2a}$$

$$S_i = \gamma Z_i + v_i \tag{2b}$$

where X_i is a set of controlling variables such as experience, gender, region, race, and economic sector of individual i. Apart from an individual's attributes such as age, gender, ethnicity, and region, the existing literature often makes use of family background, proximity to school, quarter of birth, and composition of siblings as schooling determinants (Z_i).

The OLS estimation for equation (2a) can only give a consistent estimate of λ if u_i and v_i are uncorrelated. There are many reasons why the unobserved determinants of education (u_i) and unobserved determinants of earnings or earnings residual (v_i) are correlated. For instance, individual ability may affect both an individual's education and earnings. The correlation between two disturbance terms causes ability biases in estimates of returns to schooling.

To deal with the endogeneity of education attainment and ability bias by the OLS estimator, one may take advantages of exogenous determinants of schooling decision (IV method) or compare earnings between genetically identical twins or highly genetic siblings conditional on their education attainment (within-family fixed effect) or utilize panel data. Specifically, there are four main approaches to deal with ability bias (see Belzil, 2007; Card, 2001, Griliches, 1977 for extensive surveys of the literature): *First*, employing some indicators to proxy for unmeasured ability e.g. IQ and other test scores. Because earning is positively influenced by ability so OLS estimator often provides upward-biased estimates of return to education. That is, not all of income comes from education, but a part is due to individual ability. However, ability is also affected by education, thus adding the ability proxies not only captures the effect of ability but also bias estimated returns downward (Ashenfelter, et al 1999, p. 3).

Second, using data of *siblings* or *twins*, who share the same family background and peer influences, to eliminate omitted ability bias by estimating return to schooling from difference in education attainment between siblings or twins (Ashenfelter & Zimmerman, 1997; Miller et al, 1998, Ashenfelter & Rouse, 1998, Isacsson, 1999). This strategy uses observations from the same family (twins or siblings who often have similar ability and also share the same family economic conditions) to difference out the correlation between u_i and v_i (or ability). After eliminating ability bias, the difference in earnings between siblings or twins will be attributed to difference in education among them but not due to ability. It is worth noting that, as discussed in the first approach, ability and schooling mutually affect one another, hence this approach may not provide less biased than OLS (Ashenfelter, Harmon, & Oosterbeek, 1999). Furthermore, measuring education may suffer measurement errors, level of measurement error will increase by forming differences between siblings or twins, this leads to downward biased estimates when using within-twin (or sibling) estimations (Ashenfelter, et al., 1999, p. 4).

The *third* approach is to exploit factors affecting schooling decision so as to provide instruments for schooling that are not correlated with error term of the wage equation (2a). One has to find a set of variables that affect education attainment but not earnings, this approach is called instrumental variable (IV) method. Instruments should be determinants of schooling decision, but uncorrelated with earnings residual (error term). The purpose of this method is to eliminate the differences in individual attributes between treatment group (who received more education) and control group (who received less education). Instrumental variable approach will provide a consistent estimate of the return to education (Ashenfelter, et al., 1999, p. 5). IV method first estimates effect of instrumental variables (Z) on schooling (S), then estimates the effect of the schooling (S) on earnings (y). By this procedure, the instruments affect earnings only through schooling. However, if Z_i are also correlated with earnings residual, the estimates will be biased (Angrist, Imbems, & Rubin, 1996; Staiger & Stock, 1997), especially if Z_i are weakly correlated with schooling S_i (treatment participation) and positively correlated with earnings, the estimates would be highly upward biased (Murray, 2006; Stock, 2010; Stock & Yogo, 2002). The lower the correlation between the instruments and treatment participation, the more sensitive the IV estimate is to violations of the exclusion restriction assumption (Angrist, Imbems, & Rubin, 1996, p. 451).

Another approach to overcome the ability bias is to use panel data or repeated observations over time to difference out the correlation between u_i and v_i (or ability bias). This approach exploits the variation in education and earnings over time to eliminate ability bias. This approach was initiated by Hausman and Taylor (1981), and currently applied by Arcand, d'Homebres and Gyselinck (2004), Chatelain and Ralf (2010) and others. The error term (u_{it}) of equation (2a) can be decomposed into two parts: time-invariant component (α_i) that differs across individuals and individual fixed effect (η_{it}) that is independent of both time and individuals. The panel data enable to cancel out the unobserved individual effects, ability, (α_i). Thus, the rate of return to education (λ) can be consistently estimated by the fixed effect estimator.

Fixed effect (FE) estimates return to education based on panel data for a subsample of individuals who are both working and studying over the same period. However, this approach receives several critiques. Card (1995, 1999) argues that the subsample is more likely to include individuals from poorer family background since they begin to work with low levels of education. They may pursue either full-time studying and part-time jobs or part-time studying (e.g. evening classes) and full-time jobs at a time. They lack the funds in either case to concentrate on only studying. Moreover, the model requires variation in education over time, that is, an individual has not finished schooling while they are working. This leads to another argument that the individual has a part-time or "dead-end" job, while they are studying (Card, 1994, p. 7). Additionally, these people may recognize the higher returns to schooling and decide to attain more education. Thus, one may expect that estimated return to education based on the fixed effect model is higher than the OLS estimate based on the entire sample, and so the FE estimate may not be representative of the return to education for the entire sample.

In Vietnam, on the other hand, people often have full-time job and participate in part-time classes (called "in-service training") that are often considered "very low-quality training" or diploma mill, especially higher education levels.[4] Applying fixed effect estimator may provide lower estimated return to schooling for the four-year university education than the corresponding OLS. Another limitation of fixed effect estimator is that measurement error in schooling is likely to be higher than cross-sectional estimator (Ashenfelter & Zimmerman, 1997; Belzil, 2007). Additionally, FE estimates are notoriously susceptible to attenuation bias from measurement error since measurement error often changes year to year, and often increases year to year (Angrist & Pischke, 2009). Therefore, there is more measurement error in differenced regressors of FE equation than in regressors in cross section equations.

Which is the best estimator? Comparing between alternative estimators

In IV model, the population is divided into subgroups (g) who share the same values for unobserved ability. Suppose an intervention that leads to a change (ΔS_g) in mean schooling of group g, and let β_g is the marginal return to education of group g when there is no intervention. Suppose the intervention affects only treatment group who are identical to those in comparison group, that is, they have the same unobserved ability would have the same schooling and earnings in the absence of the intervention.[5]

$$\text{plim}\beta_{iv} = E[\log y_i \mid Z_i = 1] – E[\log y_i \mid Z_i = 0]/E[S_i \mid Z_i = 1] – E[S_i \mid Z_i = 0] = E[\beta_g \Delta S_g]/E[\Delta S_g] \qquad (3)$$

If $\text{plimb}_{iv} = E[\beta_g \Delta S_g]/E[\Delta S_g] = E[\beta_g].[\Delta S_g]/E[\Delta S_g] = \bar{\beta}$ (average marginal return to education), that is, $\beta_g = \bar{\beta}$, or identical marginal return to education for all groups. However, if there exists heterogeneity in the distribution of marginal returns to school, the IV estimate based on the intervention that affects only some groups of the population will be higher or lower the

[4] see at http://thethao.tuoitre.vn/Hau-truong/415555/He-tai-chuc-da-bi-bien-tuong.html; http://www.congluan.vn/Item/VN/Thoisu/-Kem-chat-luong-do-bi-tha-noi/581E79E19ECDC140/).

[5] Suppose both groups have the same budget conditions.

corresponding OLS estimate for the same sample population. Therefore, the IV estimator in this case is referred to a Local Average Treatment Effect (LATE) since it estimates the return for subgroups who are affected by the intervention (instrument Z_i) (Imbens & Angrist, 1994). As a result, the difference between OLS and IV estimates may be attributed to difference in sample (the population vs. subgroup). The validity of IV estimator relies heavily on an assumption that the Z_i are uncorrelated with other unobserved attributes of individuals that affects earnings, that is, $cov(Z_i, u_i)=0$. In the case of experiment of Z_i (random assignment of the treatment), the difference in mean earnings between treatment and control groups will not be exacerbated by IV estimator, but this is not the case of quasi or natural experiments (Card, 1999, p. 1821). This problem is the limitation of IV estimator in estimating the returns to schooling. The IV estimates in the presence of weak instruments that are weakly correlated with schooling but have possible correlation with the residual of the earnings equation, the estimates may be very imprecise and seriously inconsistent (Belzil, 2007). Thus, weak IV test is needed to ensure that IV estimation does not provide imprecise estimates of return to education.

Difference in marginal return to education may be used to explain the difference between OLS and IV estimates. Heterogeneity of treatment effects on sub-samples can be the reason; an intervention that affects individuals with lower level of education can lead to higher IV estimates of the return to schooling relative to the OLS estimates (Card, 1994, p. 20). Therefore, programs that help improve education of children from poorer family background will tend to have higher marginal returns. Using the tuition rates and college proximity as instruments for schooling, Kane and Rouse (1993) confirm this fact.

Griliches (1977) believes that OLS estimates of education returns are unbiased or even downward biased. Similarly, according to Card (1994, 1999), the IV method yields larger estimates than the OLS. The IV studies claimed that OLS understates the returns by simply comparing wages between more and less educated workers. The difference between IV estimates and the OLS depends on the extent that instruments affect schooling decision at various levels of education due to heterogeneous returns to schooling (Card, 1999, p. 3). For example, IV estimates based on instruments which 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. In a circumstance that schooling decisions are restrained by family budget and schooling is not free of charge, instruments such as parental education and income primarily influence schooling decisions. In Vietnam, investment in education, especially higher education where there are higher costs, is constrained by household budget (Glewwe & Jacoby, 2004), so this is a good reason to believe that IV estimates may be then higher than OLS estimates for higher education since family background may strongly affect children's education attainment.

Card (1994, 1999) claims that OLS estimates of the return to education are likely to be biased downward relative to IV estimates that account for the unobserved determinants of education and earnings. Many studies reveal this typical feature. For example, Angrist and Krueger (1991) use quarter of birth as an instrument and find that IV estimate is 28% above the corresponding OLS estimate. Angrist and Krueger (1992) use the lottery numbers assigned as an instrument and find that IV estimate is 10% higher than the corresponding OLS estimate; Kane and Rouse (1993) utilize distance to colleges and tuition rates as instruments and show that IV estimate of the return to education is about 13-50% higher than the OLS estimates. Even Butcher and Case (1994) find that IV estimate is much higher than the OLS estimate (100% above the OLS) when using the presence of sisters in family as an instrument for women's schooling.

Card (1994, 1999) assumes that attenuation bias in OLS estimates of the return to education is 10-15%, hence IV estimates should exceed the corresponding OLS estimates about 10-15%, and all of the empirical studies in his survey show that the IV estimates are at least 10% above the corresponding OLS estimates. Therefore, one can conclude that cross-sectional OLS estimates of the return to schooling are biased downward relative to IV estimates which control for endogeneity of education. Furthermore, the differences in estimates of the return to education by alternative estimators may be due to measurement error in schooling (Card, 1994, p. 24). The measurement error may lead to a 10% downward bias in the OLS estimates since the OLS is based on potentially noisy measure of schooling (Card, 1994). However, Ashenfelter et al (1999) caution that the precision may be lost when moving away from the OLS estimator because the estimates based on IV estimator have larger standard errors.

IV estimator is often claimed to be able to provide less prone to mis-specification than FE estimator (Belzil, 2007; Keane, 2010). Additionally, FE estimates are often lower than both the OLS and IV estimates; this may be caused by higher measurement error from schooling measures in panel data (Belzil, 2007) since the fixed effect is highly sensitive to measurement error in schooling (Ashenfelter & Zimmerman, 1997). In summary, IV is preferred to the OLS and FE estimators when estimating returns to education, but one should bear in mind that IV estimates may be representative for sub-samples which provide a local average treatment effect (Imbens & Angrist,

1994) and precision of the estimates may be lost if the standard errors are significantly larger in comparison with that of the OLS (Ashenfelter et al, 1999).

3. Empirical models and data

Cross-sectional correlation between schooling and earnings may not reflect the true causal effect of education on earnings. One of the typical solutions to the problem of causal inference is to apply IV method as discussed in the previous section. What one needs to do is to search for instruments that affect only schooling choices but not earnings. In reality, there are two groups of IVs that belong to either supply side or demand side of schooling decision. On the supply side, many studies make use of institutional sources of schooling variation such as minimum school leaving age (Harmon & Walker, 1995), proximity to school (Card, 1995; Kane & Rouse, 1993). On the demand side, variables such as quarter of birth (Angrist & Krueger, 1991; Staiger & Stock, 1997), and family background such as parental education, year of birth, brother's education, sibling composition (Ashenfelter & Zimmerman, 1997; Butcher & Case, 1994; Card, 1995, 1999; Conneely & Uusitalo, 1997; Staiger & Stock, 1997). Hogan and Rigobon (2010) use both sides to exploit the heterogeneity in education attainment caused by differences between regions resulting from different population density, variation in the proximity to school, parental income, and income distribution, demographics, school quality, and weather etc across regions.

In our case, we look at return to the four-year university education (university graduates) using demand side factors (family background) such as parental education, assets and share of the university and post-graduated members in family as instruments. Family information such as parental education is often utilized to either directly control for unmeasured ability or as an instrument for children's schooling (Ashenfelter & Zimmerman, 1997; Card, 1995; Conneely & Uusitalo, 1997; Heckman & Li, 2004; Griliches, 1979). This is because children's education is highly correlated with their parents' characteristics especially education and economic conditions (income and assets). Card (1999, p. 1822) indicates that the correlation coefficient of parental education and children's education is about 0.4, and about 30% of the variation in US adults' education is explained by parental education. Further, we utilize household assets and parental education to proxy for permanent household income (Musgrove, 1979) which is believed to be correlated with children's education since investment in education in Vietnam is not free of charge (Glewwe & Jacoby, 2004; Glewwe & Patrinos, 1999).

Models of family background controls in return to education estimation can be set up as follows:

OLS model: $\text{Log } y_i = \alpha + \beta S_i + \lambda X_i + \delta . Z_i + \varepsilon_i$ (4)

IV model: $\text{Log } y_i = \alpha + \bar{\beta} S_i + \lambda X_i + u_i$ and $S_i = \gamma Z_i + v_i$ (5)

where $Z_i u_i$ are independent or $\text{cov}(Z_i, u_i)=0$, and $\text{cov}(Z_i, S_i) \neq 0$ or $E[Z_i, S_i] \neq 0$. S in a 0/1 variable equal to one if an individual has a bachelor's degree (the four-year university graduate) and 0 if an individual has a high school diploma. We rule out post-graduate degree holders, three-year-college and vocational-diploma holders, and below-high-school educated individuals. X_i is a set of controlling variables such as experience, experience squared, gender, ethnicity, urban, economic sectors, and eight geographical regions in Vietnam. The estimated coefficient β in equation 4 and 5 reflects a percentage difference in earnings between individuals with a bachelor's degree and high-school graduation degree. This coefficient is referred as the four-year university premium. Z_i is a set of family background such as mother's education, father's education, share of the four-year university and post-graduated members in family, and household assets (durable, fixed assets and houses) which was acquired at least one year prior to the survey.[6]

Family background (Z_i), which may be correlated with individual ability and motivation, can also be used to check the robustness of estimates by OLS estimator (Yakusheva, 2010). Even though family background variables may not be legitimate instruments for education, controlling for these variables may reduce the bias in estimated return to schooling (Card, 1999). In a review of many studies that controlled for family background, ethnicity, region and age which explain up to about 0.30 of the variance of observed schooling, Card (1994) shows that expected attenuation of the education coefficient (reduction in estimated coefficient) could be as high as 15%, this is almost as of the attenuation bias by measurement error in measured schooling. Thus, controlling for these variables also is as important as correcting for measurement error in reported education.

The most difficult task of evaluation of treatment effect is that we do not have sufficient information about subjects, people look alike but they make different choices. One does not know why people decide to take the four-year university education conditional on observed characteristics. The difference in outcomes would be affected by observed, unobserved attributes and the treatment. To measure how much they earn or return to the four-year university education, one should measure how much they would have earned if they did not have the four-year university degree (Heckman &

[6] This is to avoid the reversal causality effect of current earnings on the assets

Li, 2004). One is unable to measure the later earnings (counterfactual earnings). The conventional methods of estimating return to education do not account for the factors affecting the four-year university schooling decision, and especially investment in education in Vietnam is faced with liquidity constraints (Glewwe & Jacoby, 2004; Glewwe & Patrinos, 1999). Furthermore, the four-year university entry is not free of competition due to the government limit of quantity and facility and human capacity of education providers (universities). About one fourth of 1.2 million high school leavers are able to go to university in 2009.[7&8] Therefore, entering the four-year university education is a selectivity process.

Given the fact that to enter universities, candidates have to complete high schools and take entrance examinations, factors such as individual ability, and family resources and parental motivation play important roles in entering university education. These factors can be reflected through family background since individuals are more likely to have similar innate ability and family background than randomly selected (Ashenfelter & Zimmerman, 1997). On the supply side of the four-year university education, some studies use proximity to college as an instrument to predict schooling in Vietnam (e.g. Arcand, d'Hombres, & Gyselinck, 2004). We do not use this information since the data of distance to schools from each household measured in the current studied survey do not properly reflect the distance to school when surveyed individuals were at ages for the four-year university entry given the fact that there is a high rate of migration in Vietnam since the economic reform and almost wage-earners often reside in highly migrated regions (International Organization of Migration;[9] GSO, 2010).

Data used in this study come from Vietnam Household Living Standards Survey conducted by the Vietnam General Statistics Office in 2008 (VHLSS, 2008). The survey interviewed 9,186 households that consist of about 40,000 members covering all provinces and regions of Vietnam. The survey is representative for national level of Vietnam. From this data, we obtain 651 individuals who have either high school degree or the four-year university degree to estimate return to the four-year university education.

[7] see at http://www.business-in-asia.com/vietnam/education_system_in_vietnam.html; and http://www.gso.gov.vn/default_en.aspx?tabid=474&idmid=3&ItemID=10220

[8] Each year, the Vietnamese government allocates a certain quota of student intakes for each university depending on their facility and staff capacity.

[9] http://www.iom.int.vn/joomla/index.php

4. **Estimation results**

In this section we in sequence estimate OLS with basic controls, then with further controls of family background such as father's education, mother's education, share of university and post-graduated members in family, and household assets in logarithm. After that IV model estimation and IV tests will be conducted. Finally, Treatment Effect model estimation will be run to corroborate the IV estimates.

Unconditional wage gap between the four-year university wage-earners and high school graduated wage-earners is very large, about double (Table 1). The university graduates are more likely to work in state sector but less likely to work in private sector. They also have better family background such as parental education, assets, and have more siblings at university and post-graduated education levels. They are observed to be more in major ethnicity (Kinh and Chinese) and living in urban areas. The university graduated wage-earners are about 3 years older but have about one year of experience less than the high school graduated wage-earners (Table 1). These differences suggest either controlling for the family background variables in the OLS wage equation or using them as instruments for schooling. Additionally, when adding these variables, in sequence, into the Probit model to predict the likelihood of taking university education, we observe significant effects of these variables, that is, they meet the "*relevant*" condition, that is, $cov(Z_{i,} S) \neq 0$ (see Table 2). But when all the family background variables are added together into the model, father and mother's education turn out to be insignificant due to their highly correlation with the share of university and post-graduated members in family (the last column of Table 2). This also suggests utilizing either of them as an IV at a time.

Table 1: Summary statistics

Variables	High school graduates (n=360)		The four-year university graduates (n=291)		*t*-value for equal mean
	Mean	Std. error	Mean	Std. error	
Hourly wage (VND 1,000)	9.146	0.478	17.571	0.696	9.98**
Log of hourly wage	1.956	0.036	2.685	0.036	14.36**
Worked in state sector	0.222	0.022	0.704	0.027	13.93**
Worked in foreign sector	0.114	0.017	0.082	0.016	1.35
Worked in private sector	0.664	0.025	0.213	0.024	13.02**
Age (year)	26.706	0.288	29.793	0.368	6.61**
Experience (year)	8.706	0.288	7.801	0.367	1.94+
Gender (male=1)	0.597	0.026	0.550	0.029	1.21
Majority (Kinh & Chinese=1)	0.922	0.014	0.979	0.008	3.48**
Urban (yes=1)	0.339	0.025	0.718	0.026	10.43**
Region 1-Red River	0.286	0.024	0.268	0.026	0.51
Region 2-North East	0.097	0.016	0.107	0.018	0.39
Region 3-North West	0.039	0.010	0.014	0.007	2.05*
Region 4-North Central	0.050	0.012	0.058	0.014	0.47
Region 5-South Central	0.114	0.017	0.117	0.019	0.12
Region 6-Central Highlands	0.025	0.008	0.027	0.010	0.20
Region 7-South East	0.217	0.022	0.271	0.026	1.61
Region 8-Mekong Delta	0.172	0.020	0.137	0.020	1.22
Instruments					
Mother's education (year)	5.778	0.236	9.646	0.327	9.59**
Father's education (year)	6.331	0.264	9.405	0.385	6.58**
Share of university and post-graduated members	0.017	0.004	0.432	0.013	30.58**
Log total assets acquired before 2007	12.599	0.063	13.606	0.062	11.32**

Notes: t-value statistically significant at 10% (+), 5% (*), and 1%

Table 2: Probability of going to university

Variables	(1)	(2)	(3)	(4)	(5)	(6)
Age	0.1131	0.1113	0.0660	0.1303	0.1152	0.0736
	(3.27)**	(3.10)**	(1.55)	(3.63)**	(3.41)**	(1.71)+
Age squared	-0.0014	-0.0012	-0.0008	-0.0015	-0.0015	-0.0010
	(2.67)**	(2.26)*	(1.31)	(2.83)**	(2.84)**	(1.57)
Gender (male=1)	-0.0093	-0.0095	0.0732	-0.0405	-0.0227	0.0465
	(0.18)	(0.18)	(0.93)	(0.77)	(0.44)	(0.57)
Majority	0.1888	0.1837	0.0892	0.1908	0.0644	0.0045
	(1.41)	(1.31)	(0.73)	(1.51)	(0.46)	(0.04)
Urban	0.3422	0.2436	-0.0026	0.3064	0.1904	-0.0749
	(6.87)**	(4.39)**	(0.03)	(5.79)**	(3.18)**	(1.00)
Mother's education		0.0465				-0.0152
		(8.43)**				(1.59)
Share of university			4.0215			4.1493
and post-graduated members			(7.03)**			(7.37)**
Father's education				0.0296		0.0096
				(6.63)**		(0.87)
Log total assets					0.1880	0.1116
					(6.18)**	(2.68)**
Region dummies controlled	Yes	Yes	Yes	Yes	Yes	Yes
Wald χ^2	104.77	146.01	73.20	131.52	145.85	114.09
Prob > χ^2	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Pseudo R^2	0.1582	0.2624	0.7290	0.2208	0.2274	0.7426
Observations	651	651	651	651	651	651

*Robust z statistics in parentheses; + significant at 10%; * significant at 5%; ** significant at 1%*

OLS estimates

Estimates of return to education using OLS estimator show that university graduated wage-earners earn 71% higher than the high school wage-earners, equivalent to 17.8% per year (Table 3). When the family background is further controlled for, the return slightly declines. Interestingly, only father's education and household assets have direct effects on individual earnings, while mother's education and the share of university and post-graduated members in family have do not have such effects on earnings. This sheds some light on the validity of mother's education, father's education, the share of university and post-graduated members in family, and assets when used as IVs. We will come back to the test of IV validity later.

Table 3: Return to schooling using OLS with and without family background controls

Variables	(1)	(2)	(3)	(4)	(5)	(6)
University education	0.7134	0.6954	0.7116	0.6927	0.6561	0.6738
	(11.11)**	(10.65)**	(8.21)**	(10.72)**	(10.39)**	(7.96)**
Experience	0.0490	0.0500	0.0490	0.0550	0.0495	0.0535
	(4.11)**	(4.21)**	(4.12)**	(4.66)**	(4.27)**	(4.71)**
Experience squared	-0.0012	-0.0012	-0.0012	-0.0013	-0.0012	-0.0013
	(2.99)**	(2.94)**	(2.98)**	(3.25)**	(3.16)**	(3.35)**
Gender	0.2386	0.2379	0.2387	0.2264	0.2274	0.2189
	(4.37)**	(4.35)**	(4.38)**	(4.15)**	(4.32)**	(4.12)**
Majority	0.5764	0.5743	0.5764	0.5784	0.4941	0.5013
	(2.44)*	(2.40)*	(2.43)*	(2.46)*	(2.05)*	(2.09)*
Urban	0.0935	0.0816	0.0932	0.0813	-0.0139	-0.0114
	(1.79)+	(1.56)	(1.77)+	(1.57)	(0.24)	(0.20)
State sector	0.1351	0.1320	0.1350	0.1252	0.0846	0.0817
	(2.00)*	(1.93)+	(1.99)*	(1.85)+	(1.31)	(1.25)
Foreign sector	0.3243	0.3162	0.3243	0.3143	0.2846	0.2791
	(3.78)**	(3.56)**	(3.78)**	(3.67)**	(3.42)**	(3.26)**
Mother's education		0.0061				0.0010
		(1.06)				(0.15)
Share of university and post-graduated members			0.0050			-0.0843
			(0.03)			(0.54)
Father's education				0.0113		0.0078
				(2.47)*		(1.67)+
Log total assets					0.1242	0.1155
					(4.51)**	(4.20)**
Constant	0.6217	0.5832	0.6219	0.5018	-0.8156	-0.8080
	(3.05)**	(2.88)**	(3.04)**	(2.44)*	(2.14)*	(2.16)*
Region dummies controlled	Yes	Yes	Yes	Yes	Yes	Yes
Observations	651	651	651	651	651	651
R-squared	0.48	0.48	0.48	0.49	0.50	0.50
F-value	30.73	29.23	29.88	30.24	30.22	27.24
Prob>F	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

*Robust t statistics in parentheses; + significant at 10%; * significant at 5%; ** significant at 1%*
Note: private sector is set as a comparison base group for state and foreign sector

IV estimates

We utilized the Maximum Likelihood IV estimation (a jointly estimation procedure) and the estimates of return to education are presented in Table 4. Before presenting the results, we discuss the IV tests. The test results are presented in the bottom panel of Table 4. We emphasize the tests for the exclusion restriction or overidentification assumption and weak identification.

Table 4: Return to schooling using IV estimator with various sets of excluded instruments (LIML estimation)

Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)
University education	1.0661 (4.10)**	0.6819 (9.16)**	2.0061 (3.25)**	1.9510 (5.27)**	1.5469 (5.95)**	0.7250 (9.41)**	0.6780 (9.09)**
Experience (year)	0.0693 (3.89)**	0.0495 (4.24)**	0.1179 (2.99)**	0.1150 (4.28)**	0.0942 (4.59)**	0.0517 (4.38)**	0.0493 (4.22)**
Experience Square	-0.0017 (2.87)**	-0.0012 (2.82)**	-0.0032 (2.39)*	-0.0031 (3.25)**	-0.0025 (3.32)**	-0.0012 (2.93)**	-0.0012 (2.80)**
Constant	0.6951 (3.45)**	0.7337 (3.97)**	0.6006 (2.24)*	0.6061 (2.34)*	0.6467 (2.82)**	0.7293 (3.92)**	0.7341 (3.98)**
F-value	22.09	30.83	13.05	12.68	17.69	30.65	30.64
Prob > F	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Uncentered R^2	0.94	0.95	0.90	0.90	0.92	0.95	0.95
Root MSE	0.5841	0.5598	0.7731	0.7585	0.6626	0.5606	0.5598
Observations	651	651	651	651	651	651	651
Excluded instruments	Mother's education	Share of university and post-graduated members	Father's education	Log total assets	Mother's education & log total assets	Share of university and post-graduated members & log total assets	Share of university and post-graduated members & mother's education
Test for instruments jointly equal zero in the first stage, F-value [P-value in bracket]	29.08 [0.0000]	355.80 [0.0000]	9.97 [0.0017]	26.07 [0.0000]	22.64 [0.0000]	196.01 [0.0000]	178.19 [0.0000]
Partial R^2 of excluded instruments	0.0475	0.4837	0.0171	0.038	0.0722	0.4889	0.4843
Weak identification test (Kleibergen-Paap Wald rk F statistic) [Stock-Yogo weak id test critical value at 10% maximal LIML size in bracket]	29.08 [16.38]	355.80 [16.38]	9.97 [16.38]	26.07 [16.38]	22.64 [8.68]	196.01 [8.68]	178.19 [8.68]
Hansen J statistic (overid test) [P-value in bracket]	Just-identified	Just-identified	Just-identified	Just-identified	4.696 [0.0302]	21.735 [0.0000]	2.731 [0.0984]
Endogeity test of university education (P-val)	0.0742	0.3719	0.0019	0.0000	0.0000	0.2032	0.4240

*Robust z statistics in parentheses; + significant at 10%; * significant at 5%; ** significant at 1%. All the models controlled for gender, ethnicity, urban, economic sectors, and 8 geographical regions in Vietnam.*

First, we consider models with only one instrument at a time in columns 1, 2, 3 and 4 of Table 4. The weak identification test accepts the hypothesis that the father's education variable is a weak instrument since the Kleibergen-Paap rank F statistic (9.97) is much smaller than the Stock-Yogo's weak identification critical value at 10% maximal LIML size. Furthermore, the F-statistic on the excluded instrument in the first stage is smaller than 10. This casts doubt on the validity of the father's education as an instrument, and suggests that this instrument is weak. The point estimates are very biased and seriously inconsistent, thus, it is unable to predict the magnitude of the effects accurately when applying father's education as an instrument in IV models. In column 6, the Hansen test for exclusion restriction or over-identification rejects the validity of a combination of two instruments (the share of university and post-graduated members in family and total household assets). This implies there is at least one instrument in this combination is invalid, while in column 7 of Table 4 the combination of two instruments (the share of university and post-graduated members in family and mother's education) is accepted. This means at least one instrument is in the combination exogenous (Wooldridge, 2002).

Further, the endogeneity test in the last row of Table 4 indicates that the hypothesis of endogeneity of university education is rejected when father's education, assets, and a combination of mother and assets are used as instruments. The weak identification test statistic in column 2, which strongly rejects the hypothesis of weak instrument of the share of university and post-graduated members, and the p-value of Hansen test (column 7) is not high enough to eliminate the suspicion of a strong instrument of mother's education since power of the test is low in the presence of weak instruments, so adding a weak instrument may result in accepting the null hypothesis of overidentification just by increasing degrees of freedom (Baum, Schaffer, & Stillman, 2003). From these test results, we may infer that instruments of father's education and the household asset are invalid instruments, while mother's education is a valid but not very strong instrument, and the share of university and post-graduated members is a good instrument. This finding is contrast to Arcand, d'Hombres and Gyselinck (2004) who used a combination of father's and mother's education (parental education) as an instrument in a study of return to education in Vietnam for period 1992-1998. Mixing father's education and mother's education together may not properly reveal whose education plays an important role in schooling choices and IV modelling.

To choose which model in either column 2 or 7 of Table 4, we look at F statistic on the excluded instrument in the first stage, the F value (355.8) for the model with one variable of the share of university and post-graduated members in family doubles that (178.2) of the model with two instruments in column 7. Additionally, one should choose model among valid instruments which has a minimum mean-square error (MSE) (Donald & Newey, 2001). Furthermore, all the estimated coefficients, their standard errors, partial R^2 of excluded instruments, and MSE of these two model specifications are almost the same. This suggests that we can use either the models in columns 2 and 7 of Table 4.

Estimated return to university education varies largely. Using a weak instrument of father's education or an invalid instrument of log total household assets yields very highly upward biased results (columns 3 and 4, Table 4). Because father's education and assets are both correlated with schooling S_i (treatment participation) and positively correlated with earnings (see Tables 2 & 3), the estimates are highly upward biased (Angrist, Imbems, & Rubin, 1996; Murray, 2006; Staiger & Stock, 1997; Stock, 2010; Stock & Yogo, 2002). When using a valid instrument of mother's education the bias is reduced (comparing to estimates based on weak or invalid instruments), the return to each year of university education is about 0.27. However, when using either a strong instrument of the share of university and post-graduated members or a combination of the share of university and post-graduated members and mother's education, the return to four-year university education is 0.68, annualized return is 0.17 (columns 2 & 7). Interestingly, the estimated return using IV models with valid instruments is almost the same with that based on the OLS model with family background controls (column 6 of Table 3), even somewhat lower than the OLS estimate based on the model without family background control (column 1 of Table 3). This implies there is no a serious ability bias in the OLS estimated return to the university education in Vietnam.

Treatment Effect Model estimates

In the above IV estimation with the joint estimation procedure (treatment participation and outcome equation), the normal distribution assumption of the first stage dependent variable was ignored even though it is a binary variable. The joint estimation procedure may be acceptable since the OLS still remain unbiased (Gurajati, 1995, p. 543). However, the estimates that ignored the assumption may be woefully inefficient (Nichols, 2009).

Treatment effect model may be an alternative approach to the problem of non-fulfilment of the normality assumption of binary endogenous variable of university education in the first stage. The binary endogenous regressor of university education is viewed as a treatment indicator, thus this estimation is considered as the treatment effect model (Heckman & Li, 2004). Error terms (u_i of main equation, and v_i of instrumental equation) are assumed to be correlated, i.e. $cov(u_i, v_i) = \rho\sigma^2$

Table 5: Return to schooling using Treatment Effect Model with various sets of controlling variables in selection equations

Controls in wage equation	**(1)**	**(2)**	**(3)**
University education (yes=1)	0.7830	0.7113	0.7030
	(3.33)**	(9.27)**	(9.07)**
Experience (year)	0.0520	0.0489	0.0485
	(3.65)**	(4.17)**	(4.12)**
Experience squared	-0.0012	-0.0012	-0.0011
	(2.76)**	(3.02)**	(2.98)**
Gender (male=1)	0.2347	0.2388	0.2392
	(4.16)**	(4.42)**	(4.42)**
Majority (Kinh & Chinese=1)	0.5698	0.5766	0.5774
	(2.42)*	(2.46)*	(2.47)*
Urban (yes=1)	0.0722	0.0941	0.0967
	(0.89)	(1.80)+	(1.84)+
State sector (yes=1)	0.1018	0.1361	0.1401
	(0.77)	(1.92)+	(1.98)*
Foreign sector (yes=1)	0.3150	0.3245	0.3257
	(3.32)**	(3.82)**	(3.83)**
Constant	0.6128	0.6220	0.6230
	(3.00)**	(3.09)**	(3.10)**
Region dummies controlled	Yes	Yes	Yes
Controls in selection equation (the first stage)			
Variables as of the wage equation	Yes	Yes	Yes
Mother's education	Yes		Yes
Share of university and post-graduated members		Yes	Yes
Wald χ^2	332.57	434.96	342.01
Prob > χ^2	0.0000	0.0000	0.0000
Observations	651	651	651

*Robust z statistics in parentheses; + significant at 10%; * significant at 5%; ** significant at 1%*

Note: private sector is set as a comparison base group for state and foreign sector

where $u_i \sim NID(0, \sigma^2)$ and $v_i \sim N(0,1)$. This model offers an estimator similar to IV estimator in the case of a single binary endogenous variable, but it improves efficiency of estimates (Nichols, 2009, p. 56). For the treatment effect model, the Lambda or inverse Mills' ratio is estimated in the first stage and then is included in the second stage to correct for selection bias. The identification is obtained by including factors (as of the valid instruments in the IV models above) that influence university education participation but not earnings. The estimates are presented in Table 5. The estimated return to university education (17.8% and 17.6% per year for model with the share of university and post-members in family-column 2, and a combination of the share of university and post-members in family and mother's education-column 3, respectively) seems to accord with the estimates based on the previous IV models.

5. **Concluding remarks**

This paper utilizes a recent dataset to estimate the return to post-in Vietnam. We demonstrate that controlling for proxy for individual ability (family background) in the wage equation slightly reduces the estimated return to higher education. This trend is also true when mother's education and share of university and post-graduated members in family are used as instruments in IV models. Therefore, OLS estimates are upward-biased, but the bias is not large to be concerned. Additionally, the paper demonstrates that using invalid or weak instruments, such as father's education and household assets, leads to highly imprecise estimates of the return to the four-year university schooling.

In 2008 income premium for university education in Vietnam is about 68% above the high school education (on average, 17% per year). The return to higher education reached the average return of higher education, about 18%, in Asia (Psacharopoulos & Patrinos, 2004). The estimated return seems to be robust to various estimators of OLS, IV and Treatment Effect. The return to university education is much higher than that of ten years ago in 1998 (Doan & Gibson, 2009). This implies that labour market in Vietnam rewards higher-skilled workers more after a longer period of economic transition to a market economy. This increasing trend is also observed in a compatibly transitional economy of China (Heckman & Li, 2004). The high premium for university education may be also attributed to university graduates' comparative advantage in the Vietnam labour market where only 5% of the population hold university or post-graduate degrees (GSO, 2010).

References

Angrist, J., & Krueger, A. (1991). Does compulsory schooling attendance affect schooling and earnings?. *Quarterly Journal of Economics, 106*(4), 979-1014.

Angrist, J., & Krueger, A. (1992). *Estimating the payoff to schooling using Vietnam-era draft lottery* (Working paper no. 4067). Cambridge, MA: National Bureau of Economic Research.

Angrist, J., & Pischke, J. (2009). *Mostly harmless econometrics.* Princeton and Oxford: Princeton University Press.

Angrist, J., Imbens, G., & Rubin, D. (1996). Identification of causal effects using instrumental variables. *Journal of the American Statistical Association, 91*(434), 444-455.

Arcand, J., d'Hombres, B., & Gyselinck, P. (2004). *Instrument choice and the returns to education: New evidence from Vietnam*. Retrieved from http://129.3.20.41/eps/lab/papers/0510/0510011.pdf

Ashenfelter, O., & Rouse, C. (1998). Income, schooling and ability: evidence from a new sample of identical twins. *Quarterly Journal of Economics,* 113, 253-284.

Ashenfelter, O., & Zimmerman, D. (1997). Estimates of the return to schooling from sibling data: fathers, sons and brothers. *Review of Economics and Statistics, 79*(1), 1-9.

Ashenfelter, O., Harmon, C., & Oosterbeek, H. (1999). A review of estimates of the schooling/earnings relationship, with tests for publication bias. *Labour Economics, 6*(4), 453-470.

Baum, C., Schaffer, M., & Stillman, S. (2003): Instrumental Variables and GMM: Estimation and Testing. *Stata Journal, 3*(1), 1-31.

Belzil, C. (2007). The return to schooling in structural dynamic models: a survey. *European Economic Review, 51*(5), 1059-1105.

Butcher, K., & Case, A. (1994). The effect of sibling sex composition on women's education and earnings. *The Quarterly Journal of Economics, CIX*(3), 531-563.

Card, D. (1994). *Earnings, schooling, and ability revisited* (Working paper no. 4832). Cambridge, MA: National Bureau of Economic Research.

Card, D. (1995). Using geographic variation in college proximity to estimate the return to schooling. p.201-222, Aspects of labour market behavior: Essays in honour of John Vanderkamp.

Card, D. (1999). The casual effect of education on earnings. In *Handbook of Labor Economics*. Vol 3A, ed. By Ashenfelter, O., & Card, D. Amsterdam and New York: North Holland Publisher.

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

Chatelain, J., & Ralf, K. (2010). *Inference on time-invariant variables using panel data: A pre-test estimator with an application to the returns to schooling.* Retrieved from http://hal-paris1.archives-ouvertes.fr/docs/00/49/20/39/PDF/Chatelain_Ralf_Time_Invariant_Panel.pdf

Conneely, K., & Uusitalo, R. (1997). *Estimating heterogeneous treatment effects in the Becker schooling model* (Unpublished discussion paper). Industrial Relations Section, Princeton University.

Doan, T., & Gibson, J. (2009). *Do returns to schooling go up during transition? The not so contrary case of Vietnam* (Working papers in economics no. 09/08). Retrieved from University of Waikato, Department of Economics website: ftp://mngt.waikato.ac.nz/RePEc/wai/econwp/0908.pdf

Donald, S., & Newey, W. (2001): Choosing the Number of Instruments. *Econometrica*, 69(5), 1161-1191.

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

Glewwe, P., & Patrinos, H. (1999). The role of the private sector in education in Vietnam: Evidence from the Vietnam Living Standards Survey. *World Development, 27*(5), 887-902.

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

GSO. (2010). *Population and Housing Census 2009, 1 April 2009* (Complete report and major findings, General Statistical Office). Hanoi, Vietnam.

Gujarati, D. (1995). *Basic Econometrics.* New York, London: McGraw-Hill, Inc.

Hausman, J., & Taylor, W. (1981). Panel data and unobservable individual effects. *Econometrica, 44*(6), 1377-1398.

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

Hogan, V., & Rigobon, R. (2010, June). *Using unobserved supply shocks to estimate the returns to education.* Paper presented at the European Association of Labour Economists Conference at University College London, England. Retrieved from http://www.eale.nl/Conference2010/Programme/PaperscontributedsessionsI/add127862_ElgN5Qi516.pdf

Imbens, G., & Angrist, J. (1994). Identification and estimation of Local Average Treatment Effects. *Econometrica, 62*(2), 467-476.

Isacsson, G (1999). Estimates of the return to schooling in Sweden from a large sample of twins. *Labour Economics, 6*(1999), 471–489.

Kane, T., & Rouse, C. (1993). *Labour market returns to two- and four-year colleges: is a credit a credit and do degrees matter?* (Working paper no. 4268). Cambridge, MA: the National Bureau of Economic Research.

Keane, M. (2010). Structural vs. atheoretic approaches to econometrics. *Journal of Econometrics, 156*(1), 3-20.

Miller, P., Mulvey, C., & Martin, N. (1995). What do twins studies reveal about the economic returns to education? A comparison of Australian and U.S findings. *American Economic Review*, 85, 586-599.

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

Murray, M. (2006). Avoiding invalid instruments and coping with weak instruments. *Journal of Economic Perspectives, 20*(4), 111-132.

Musgrove, P. (1979). Permanent household income and consumption in urban South America. *American Economic Review*, 69, 355-368.

Nichols, A. (2009). *Casual inference with observational data: Regression Discontinuity and related methods in Stata*. Retrieved from http://repec.org/nasug2007/causal.pdf

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

Staiger, D., & Stock, J. (1997). Instrumental variables regression with weak instruments. *Econometrica, 65*(3), 557-586.

Stock, J. (2010). The other transformation in econometric practice: Robust tools for inference. *Journal of Economic Perspectives, 24*(2), 83-94.

Stock, J., & Yogo, M. (2005). Testing for Weak Instruments in Linear IV Regression. In Andrews, D & Stock, J (Eds), *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg* (pp. 80–108). Cambridge: Cambridge University Press.

Stock, J., & Yogo. (2002). *Testing weak instruments in linear IV research regression* (NBER Technical working paper no. 284). Retrieved from the National bureau of Economic Research website: http://www.nber.org/papers/T0284

Stock, J., Wright, J., & Yogo, M. (2002). A survey of weak instruments and weak identification in generalized method of moments. *Journal of Business & Economic Statistics, 20*(4), 518-529.

VHLSS. (2008). *Vietnam Household Living Standards Survey 2008* [Datasets and basic results]. General Statistic Office (GSO), Hanoi, Vietnam.

Wooldridge, J. (2002): *Econometric Analysis of Cross-Section and Panel Data*. Cambridge, MA: MIT Press.

Yakusheva, O. (2010). Return to college education revisited: is relevance relevant? *Economics of Education Review, 29*(2010), 1125-1142.