



Skill formation and transition to productive livelihood in Vietnam

Ngo Thi Minh Tam Tran

► To cite this version:

Ngo Thi Minh Tam Tran. Skill formation and transition to productive livelihood in Vietnam. Economics and Finance. Université Paris sciences et lettres, 2017. English. NNT : 2017PSLED033 . tel-01782405v2

HAL Id: tel-01782405

<https://theses.hal.science/tel-01782405v2>

Submitted on 14 May 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

THÈSE DE DOCTORAT

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

Préparée à l'Université Paris-Dauphine

Skill formation and transition to productive livelihood in Viet
nam

École Doctorale de Dauphine — ED 543

Spécialité Sciences économiques

COMPOSITION DU JURY :

Présidente du jury

Mme Christelle DUMAS
University of Fribourg

Rapporteur

M. Remco OOSTENDORP
VU Research Institute

Membre du jury

M. Martin Rama
The World Bank

Membre du jury

Mme Laure PASQUIER-DOUMER
IRD-DIAL

Directeur de thèse

M. François ROUBAUD
IRD-DIAL

Soutenue le 13.12.2017
par Ngo Thi Minh Tam TRAN

Dirigée par François ROUBAUD

UNIVERSITÉ PARIS DAUPHINE
ÉCOLE DOCTORALE DE DAUPHINE

Skill formation and transition to productive livelihood in Vietnam

THESE

Pour l'obtention du titre de Docteur en Sciences Économiques

Ngo Thi Minh Tam TRAN

Jury

Christelle DUMAS, University of Fribourg	Rapporteure + Présidente du jury
Remco OOSTENDORP, VU Research Institute	Rapporteur
Martin RAMA, The World Bank	Membre du jury
Laure PASQUIER-DOUMER, Institut de Recherche pour le Développement	Membre du jury
François ROUBAUD, Institut de Recherche pour le Développement	Directeur de thèse

♥ *Dành tặng Mẹ thương yêu* ♥

♥ *Tặng người bạn đời và các con yêu dấu* ♥

Table of Contents

Table of Contents	4
List of Tables.....	6
List of Figures	7
Acknowledgments.....	11
General introduction	12
1. The context of Vietnam.....	13
2. Researches on skills and educational inequalities for Vietnam	22
3. Objective, research questions and associated chapters	24
3.1. Chapter 1: Skills as the predictors of lower secondary dropping out in Vietnam: an estimation of Cox proportional hazard model.....	25
3.2. Chapter 2: Does full-day schooling reduces educational inequality in Vietnam?.....	26
3.3. Chapter 3: Does cognitive or non-cognitive skill matter in predicting labour market earnings: Evidence from Vietnam?	26
4. Skills: concepts and measurement.....	27
5. Conceptualising skills in conventional economic models	29
6. Data.....	30
6.1. Young Lives Household Surveys and School Survey in Vietnam	30
6.1.1. Young Lives Household Surveys.....	30
6.1.2. School Survey 2011-12	31
6.2. Measures of cognitive and non-cognitive skills in the Young Lives Survey.....	32
6.2.1. Cognitive skills	33
6.2.2. Non-cognitive skills.....	34
References.....	37
Appendix	41
A. Variables included in scales of non-cognitive skill	41
Chapter 1: Skills as the predictors of lower secondary dropping out in Vietnam: An estimation of Cox proportional hazard model.....	46
Abstract	46
1. Introduction	47
2. Literature review.....	49
2.1. International studies on the determinants of school dropout	49
2.2. Studies on dropping out for Vietnam	54

3. Education in Vietnam: an overview	56
3.1. Education system.....	56
3.2. Performance of the education sector	57
4. Data and methodology	58
4.1. Data.....	58
4.2. Dropouts	58
4.3. Measures of cognitive and non-cognitive skills	61
4.4. Estimation strategy	61
5. Determinants of dropping out in Vietnam: empirical results	65
6. Conclusions and policy discussion.....	77
Reference	79
Appendix	83
A. Tables	83
B. Figures.....	93
Chapter 2: Does full-day schooling reduce educational inequality in Vietnam?	97
Abstract	97
1. Introduction	98
2. Full-day schooling in the context of Vietnam.....	100
3. Data and definitions.....	101
4. Does FDS reduce the gap in instructional time between students with different social backgrounds?	104
5. Do students from different social backgrounds have the same access to full-day schooling resources?	108
6. Does full-day schooling have a positive effect on learning achievement varying by social background?	110
6.1. Estimation Strategy.....	111
6.2. Estimation Results	113
6.3. Robustness checks	117
7. Conclusion.....	119
Acknowledgements.....	121
References.....	121
Appendix	123
A. Tables	123
B. Figures.....	133

Chapter 3: The relative importance of skills in predicting labour market earnings in Vietnam.....	136
1. Introduction	137
2. Skills and labour market earnings: a literature review	139
2.1. International literature	139
2.2. Related literature for Vietnam.....	143
3. Data:.....	144
3.1. Measures of cognitive and non-cognitive skills	144
3.2. Data on employment and earnings	145
4. Estimating returns to skills: empirical strategy	147
5. Skills and labour outcomes	150
5.1. Skills and labour earnings: descriptive analysis using the Young Lives data.....	150
5.2. Skills and Education: estimation results	152
5.3. Relative importance of skills in labour outcomes	155
6. Conclusions and discussion	159
6.1. Findings.....	159
6.2. Discussion on policy implications	159
References.....	162
Appendices	167
A. Summary on the evidences on skills and labour outcomes	167
B. Tables	173
C. Figures.....	175
Résumé de la thèse	176

List of Tables

General introduction

Table 0- 1: Indices included in the composite index of cognitive and non-cognitive skill.....	32
Table 0- 2: Cronbach's alpha for the non-cognitive scales.....	36

Chapter 1

Table 1- 1: Number of dropouts by highest grade completed and academic year	59
Table 1- 2: Variables included in the models	63
Table 1- 3: Description of time-constant variables	66

Table 1- 4: Description of time varying variables	67
Table 1- 5: Cox models with single and composite measures of cognitive and non-cognitive skills....	69
Table 1- 6: Cox models with single cognitive and non-cognitive skills	70
Table 1- 7: Cox models with composite cognitive and non-cognitive skills	72
Table 1- 8: Baseline survival and hazard of selected groups	75

Chapter 2

Table 2- 1: Average social background indices for students with selected factors	104
Table 2- 2: Instructional time and access to FDS and extra classes by quintile of social background	105
Table 2- 3: Access to SEQAP schools by social background (percentage)	105
Table 2- 4: Mean instructional time across schooling types, the entire sample and student social backgrounds (number of periods)	107
Table 2- 5: Number of schooling periods by subject by FDS status and social background	107
Table 2- 6: Frequency of students with school resources – FDS vs. non-FDS (%)	108
Table 2- 7: Estimation results of learning progress with 2SLS (Beta coefficients)	115
Table 2- 8: Estimation results of learning progress with 2SLS by distinguishing between FDS with and without SEQAP (Beta coefficients)	116
Table 2- 9: Estimation results of learning progress with 2SLS with alternative measures of FDS (Beta coefficients)	118
Table 2- 10: Estimation results of learning progress with OLS on the entire School Survey sample (Beta coefficients)	119

Chapter 3

Table 3- 1: Pairwise correlation between skills (age 19)	144
Table 3- 2: Characteristics of the sample	146
Table 3- 3: Skills and earnings, by group	150
Table 3- 4: Model of cognitive and non-cognitive skills	153
Table 3- 5: Model of cognitive and non-cognitive skills: Choices on educational level	154
Table 3- 6: Model of cognitive and non-cognitive skills: labour market participation and employment choice	156
Table 3- 7: Model of cognitive and non-cognitive skills: Log of hourly earnings, whole sample	158

List of Figures

General introduction

Figure 0- 1: Net enrolment rate, primary and lower secondary education (percentage)	13
---	----

Figure 0- 2: Educational attainment across country income level	14
Figure 0- 3: Changes in net enrolment rate, by educational and income (percentage).....	16
Figure 0- 4: Change in net enrolment rate, by education and ethnicity (percentage).....	16
Figure 0- 5: Change in net enrolment rate, by education, gender and region (percentage)	17
Figure 0- 6: Mean score (percentile) - in Maths and Vietnamese by gender, wealth and parents' education	17
Figure 0- 7: Dropout rate by educational level completed (percentage).....	18
Figure 0- 8: Non-cognitive achievement of Vietnam compared to other developing countries'	19
Figure 0- 9: Attendance in private tutoring, by education (percentage)	21
Figure 0- 10: Thesis in a nutshell	24

Chapter 1

Figure 1- 1: Vietnam's education system	56
Figure 1- 2: Dropout rates, by highest educational level completed (%)	58
Figure 1- 3: Dropout rates, by highest lower secondary grade/level completed (%).....	60
Figure 1- 4: Dropout rates, by highest lower secondary grade completed and academic year (%).....	60
Figure 1- 5: Baseline cumulative hazard and hazard contribution	68
Figure 1- 6: Cumulative hazards and hazard contribution, subgroups with skills equal to dropout's mean skills	74
Figure 1- 7: Baseline cumulative hazards, subgroups with skills less than dropout's mean skills.....	75
Figure 1- 8: Estimated hazard functions, whole sample and sub-group	76

Chapter 2

Figure 2- 1: Distribution of student social backgrounds by schooling type.....	106
Figure 2- 2: School resources by social background level for students attending FDS	110

Chapter 3

Figure 3- 1: Distribution of skill indexes by level of education.....	145
Figure 3- 2: Sample structure	146
Figure 3- 3: Distribution of skills, by choice on labour market participation	151
Figure 3- 4: Distribution of skills, by employment type	152
Figure 3- 5: Distribution of estimated latent cognitive and non-cognitive skill by education.....	155
Figure 3- 6: Skills and employment choices	157
Figure 3- 7: Skills and earnings	158

List of Appendices

General introduction

A. Variables included in scales of non-cognitive skill	41
--	----

Chapter 1

A. Tables.....	83
A.1. Pairwise correlation between explanatory variables.....	83
A.2. Factor analysis for the composite measure of the cognitive skill	84
A.3. Test for the significance of variables included in the model.....	85
A.4. Test for proportional hazard assumption based on Schoenfeld residuals.....	86
A.5. Test for proportional hazard assumption for Models (14) and (21)	87
A.6. Robustness check – Cox models with single measures of cognitive skill	89
A.7. Robustness check – Cox models with composite measures of cognitive skill.....	91
B. Figures.....	93
B.1. Goodness of fit of the full model (7) – individual measures of skills	93
B.2. Goodness of fit of the full model (7)– composite measures of skills	93
B.3. Estimated survival function.....	94
B.4. Estimated cumulative hazard, whole sample and subgroups.....	95

Chapter 2

A. Tables.....	123
A.1. Descriptions of variables used for the social background index	123
A.2. MCA for Social background index	124
A.3. MCA Statistics of columns – Social background index	125
A.4. MCA for School facilities index.....	126
A.5. MCA Statistics of columns - School facilities index	126
A.6. School facilities index by facility.....	127
A.7. MCA for Class facilities index	127
A.8. MCA Statistics of columns - Class facilities index.....	128
A.9. Class facilities index by facility	129
A.10. Description of variables	129
A.11. Selection equation estimation results	130
A.12. Whole estimation results of learning progress in Mathematics (2SLS)	131
A.13. Whole estimation results of learning progress in Vietnamese (2SLS)	132

B.	Figures.....	133
B.1.	Distribution of students across the social background index	133
B.2.	Statutory teaching hours, 2003.....	133
B.3.	Statutory teaching hours, 2012.....	134

Chapter 3

A.	Summary on the evidences on skills and labour outcomes.....	167
B.	Tables.....	173
B.1.	Factor analysis for cognitive and non-cognitive skill index	173
B.2.	List of explanatory variables included in outcome equations.....	174
B.3.	Tests for normality of estimated latent skills.....	174
C.	Figures.....	175
C.1.	Distribution of skills.....	175

Acknowledgments

I would like to express my sincere gratitude to my supervisors, François Roubaud and Laure Pasquier-Doumer for their most ceaseless supports and favours. They have devoted their essential guidance and valuable advice to my thesis. Their thoughtfulness and earnest manner have stimulated my hard working and nurtured my desire for obtaining research papers in their best states. May I extend my heartfelt thanks to my dear mentor Laure Pasquier-Doumer who is the co-author on my Chapter 2. She has accompanied me since the embryo of my PhD project and provided me with sharp remarks on my research's rough idea. Without her kind-hearted and patient assistance, I would not have accomplished the thesis in time.

I am very grateful to Christelle Dumas and Remco Oostendorp for being my thesis reporters and Martin Rama for holding the chair in my PhD defence. Their constructive comments helped me improve my research a lot.

I owe Dr Nguyen Thang, my boss at the Centre for Analysis and Forecasting (CAF) under the Vietnam Academy of Social Sciences for giving me the opportunity to work as a researcher in the NOPOOR Project, where I granted a PhD scholarship and for his facilitation of the process of my PhD research. I would like to thank the Young Lives Project providing me with the access to the Young Lives data which is my thesis data source, and for partially funding my second chapter. I am also much indebted to Xavier Oudin, the deputy representative of the Institut de Recherche pour le Développement in Vietnam and coordinator of the European project NOPOOR for his very thoughtful support in all aspects of doing research and administration.

My sincere thanks go to Christophe Nordman (IRD-DIAL), Axel Demenet (IRD-DIAL), my colleagues at CAF, La Hai Anh and Vu Hoang Dat for their valuable exchanges, as well as other colleagues at CAF and my friends for their sharing and encouragement whenever I get stuck in my research.

I highly appreciate the friendly atmosphere in DIAL. I would like to show my wholehearted thanks to all DIAL members for their hospitality and kind supports research whenever I was in Paris for my PhD. I ask for forgiveness if I miss some of their names - Flore, Sandrine, Mireille, Danielle, Anne, Loïc, Delia, Camille, Sarah, Marlon, Estelle, Marin, Björn. I will always be thankful for the time and the kindness that my Vietnamese friends in DIAL, Nguyen Huu Chi, Hoang Thuy Quynh, Le Thi Thuy Linh, Nguyen Tuan Minh and my high school alumni in Paris extended to me.

I am thankful to the IRD for granting me the PhD scholarship through the "Allocations de recherche pour une thèse au Sud". Without this program, this thesis cannot be realised. I also acknowledge Paris-Dauphine University for their administrative support.

Last but not least, I wish to demonstrate my most profound indebtedness to my beloved late mother, my dear husband and son for their love, encouragement, moral and material support. They give me strength and energy to complete the thesis successfully.

General introduction

"Should you do your best for your schooling if you wish to be happy when growing up. Otherwise, you will not be able even to earn your living", reminded my mother a thousand times when I was a small girl. She is like most of the parents in Vietnam who always appreciate the value of schooling. Early in the feudalism, Vietnamese considered succeeding in life as passing examinations. Anyone achieving a level of education (high-school 'đỗ tú tài', bachelor level 'đỗ cử nhân' or first doctoral candidate 'đỗ trạng nguyên') brings fame to his family and ancestors. Subject to the educational level achieved, the scholars would receive greeting ceremony from the commune/district/provincial government officials and people. Until now the perception seems to hold as people often look at the educational attainment or learning achievement as signals for prosperous livelihood. More distortedly, people exaggerate the worth of test scores as the indication of successful schooling. The learning process has become examination-oriented, accompanied by the all-over private tutoring.

My personal experience does not prove that good performance at school means later good outcomes in life. There should be other factors of the following results rather than the school performance. That is why many of my friends who were unexceptional performers at school turned to become outstanding individuals with more favourable careers and more comfortable circumstances than the best performers. Moreover, it is puzzling that Vietnam de facto has achieved exceptional performance in educational attainment compared to other countries even with higher income but had the lowest labour productivity in Asian-Pacific region. Putting aside the differences in conditions in which people are born and reared, the contrary outcomes of education and livelihood might derive from some problems in the education sector. Are there obstacles to obtaining adequate skills for productive livelihood? What are they and what should be done to address them? The questions prompt my thesis. Understanding these drawbacks contributes a voice to the reforming of the education sector for more even opportunities of all in the country.

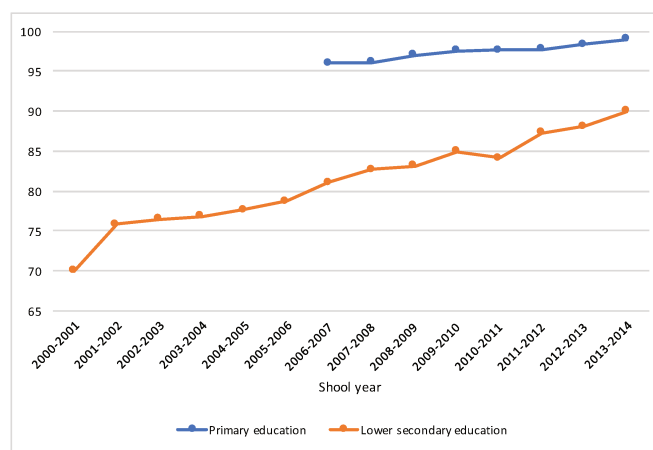
The remainder of the general introduction is structured as follows. Section 1 presents the context of Vietnam's education and its paradox. Section 2 shows ambiguous evidence that gives rise to the need for the study in this thesis. Section 3 provides the contextual information on the research questions tackled in each chapter as well as chapter summaries. Section 4 describes the data and terminologies used for the whole thesis. Finally, Section 5 supplies the definitions and measurements of the terms in theory/general and the chapters.

1. The context of Vietnam

Exceptional achievements in educational attainment and acquisition of cognitive skill

Vietnam has attained significant successes in economic development and poverty reduction for the past two decades. The growth rate of Gross Domestic Product averaged 7.2 percent per annum between 1993 and 2008.¹ Poverty rate sharply decreased from 58 to about 15 percent over the period.² Inequality has been almost stable since 2000³ given the high growth rate over the period. Alongside the broad-based economic growth, Vietnam's education sector has been performing well. Compared with the countries of same income level, its adult population has more (mean) years of schooling (Figure 0- 2A). The enrolment rates at all levels increased dramatically and much higher than that of other low-income countries (Glewwe, 2004b; Nguyen, 2004). For basic schooling, primary enrolment has been nearly universal, with the net enrolment rate of 99 percent in the school year 2013-2014 (Figure 0- 1). Meanwhile, net enrolment rates for lower secondary education continuously rose from 70 to 90 percent between 2000 and 2014.

Figure 0- 1: Net enrolment rate, primary and lower secondary education (percentage)



Source: Ministry of Education and Training

More remarkable achievement of the education sector is the outstanding student's acquisition of cognitive skill relative to even countries with the much higher level of income. Cognitive skill is defined

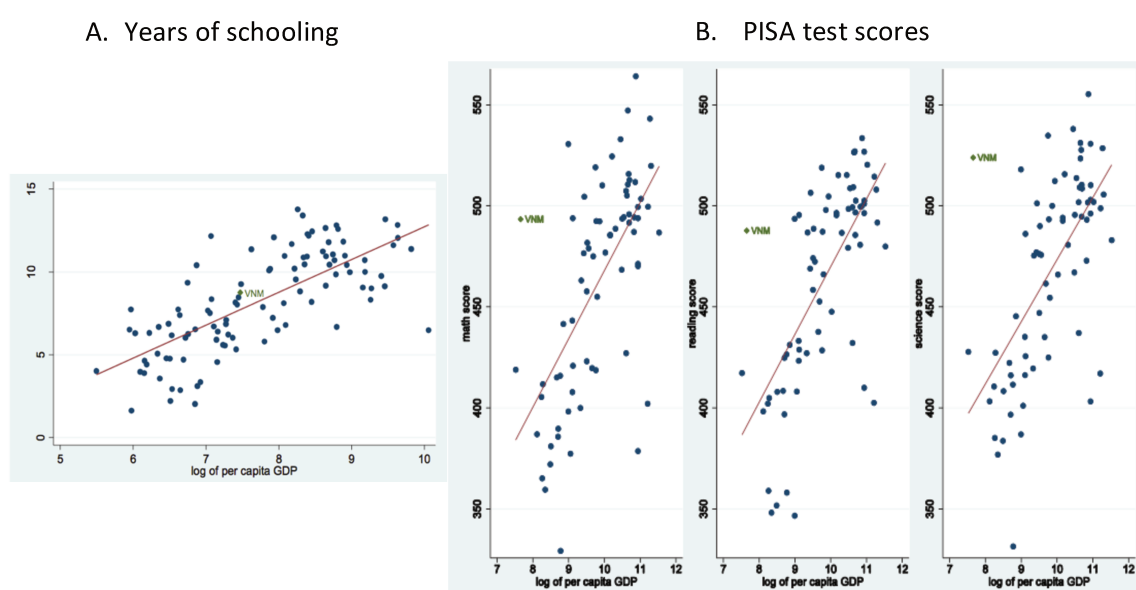
¹ Calculated based on data from the General Statistics Office (GSO)

² The poverty rates are estimated by GSO based on the Vietnam Living Standard Surveys (VLSSs) conducted by GSO in 1993, 1998, Vietnam Household Living Standard Surveys (VHLSSs) in 2002, 2004, 2008, 2010, and the poverty line set by the GSO and the World Bank.

³ Gini coefficient of income/expenditure distribution rising in the 1990s and relatively stable in the 2000s at the level of 0.43 and 0.37 on average respectively (GSO, 2012; VASS, 2010). Gini slightly declined from 0.375 in 2002 to 0.36 in 2014 (Benjamin et al., 2017).

by the ability to solve the abstract problems and measured by the cognitive test scores such as IQ or math test. In the OECD's Program for International Student Assessment (PISA) of 15-year-old student's performance in science, math and reading, Vietnam was ranked 8th, 17th and 19th out of 65 participating countries in 2012. In 2014, the corresponding figures were and 8th, 22nd and 32nd of 72 countries in 2015 (Figure 0- 2B). The rankings on all subjects put Vietnam above the OECD average as well as, excluding reading skill in 2015, above the United States, Australia, United Kingdom and France. According to the results from the STEP⁴ surveys (Roseth, Valerio, & Gutiérrez, 2016), Vietnam's labour force held the first in the literacy score and the second in the intensity of use in reading and writing skills among developing countries. Given the performance in cognitive skill, Vietnam is an outstanding example worth replicating in other countries.

Figure 0- 2: Educational attainment across country income level



Note: The green dots labelled “VNM” represent Vietnam among countries (in blue dots)

Source: extracted from Dang & Glewwe (forthcoming)

But gaps in skills and low non-cognitive skill

However, there have been drawbacks to further progress of the sector and the social and economic development. First, inequalities and inequality of opportunities in well-being and education

⁴ The World Bank's Skills Towards Employability and Productivity – Skill Measurement Program that consists of a household survey and an employer survey. The program collected information on supply and demand for skills in urban areas in 12 developing countries, including Bolivia, Colombia, Ghana, Lao PDR, Ukraine, Vietnam, and the Yunnan Province in China in 2012 and Armenia, Azerbaijan, Georgia, Kenya, and Macedonia FYR in 2013.

widen the gaps in skills. Despite little changes in inequality⁵, there remain disparities in the outcomes and opportunities between various segments of the society. Income inequality rose for the upper half of the distribution and rural regions. Ethnic differences, which were higher among minorities than among the Kinh (the majority), mainly accounted for rural income inequality. Moreover, within-location disparities - between rural and urban areas/provinces/rural and urban areas within a province explain increasingly part of the inequality (Benjamin, Brandt, & McCaig, 2017).

Also, public concern about rising inequalities and inequalities in access to education (among other aspects such as health care, job and land) is growing over time.⁶ In 2012, over a half of rural residents and 76 percent of urban residents were concerned about inequality in living standards. Notably, the rural group most worried about quality and access to health services, while urban group the quality and access to education. The inequalities are thus accompanied by the inequality of opportunities (Kozel, 2014; Oxfam, 2017; World Bank, 2014). Opportunities refer to the circumstances which children are born into, such as gender, ethnicity, parental education and income, that affect their access to basic services (e.g. education and health services) thereby affect their later outcomes in life. Evidence for Vietnam implies that the deficits in learning are attached to family background and early set before the age of formal schooling (Bodewig & Badiani-Magnusson, 2014).⁷ The share of 5-year-old children in poor households that are vulnerable (scored in the lowest decile) in at least one domain of child development – 40 percent doubled that of the non-poor children (ibid).

Together with inequalities in well-being and opportunities, there exist differences in educational participation, educational attainment and achievement across regions and levels of income/expenditure. The overall inequality in educational attainment is relatively low with a Gini coefficient of 0.25. There were large variances in provincial and ethnic Gini coefficients (from 0.18 to 0.53) and years of schooling (from 4 to 11 years), and small gender disparity (gender difference of 0.04 in Gini) (Rew, 2008).

Inequalities in access to education are low for the primary level but considerable for the lower and upper secondary level. In 2012, the net enrolment rate in lower and upper secondary level of the highest income achieved 94 and 83 percent respectively. Meanwhile, the corresponding shares for the lowest income quintile were 71 and 36 percent (Figure 0- 3). A similar trend was observed for ethnic gaps: compared to the minorities, the majorities (Kinh and Hoa group) had 10 and 30 percentage point higher in the net enrolment rate at lower and upper secondary education in 2014

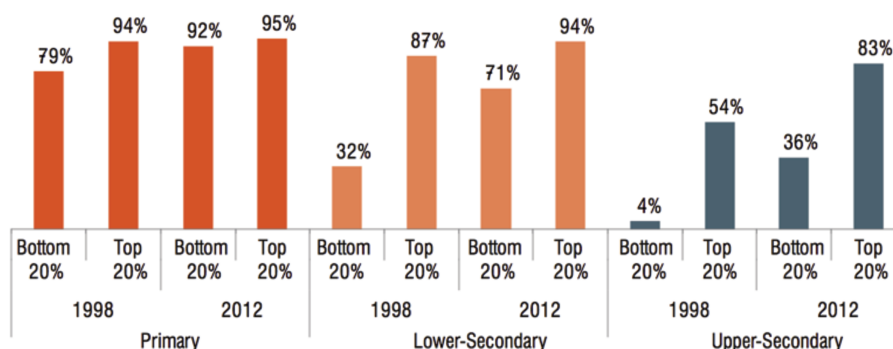
⁵ Gini slightly declined by only 0.015 from 0.375 in 2002 to 0.360 in 2014 (Benjamin et al., 2017).

⁶ According to World Bank (2014b) and Oxfam (2017) that based on the Perceptions of Inequality Survey conducted by the World and the Institute of Labour Sciences and Social Affairs Bank in 2012, and Qualitative Survey on Perception of Inequality funded by Oxfam in 2012.

⁷ According to the Ministry of Education and Training's assessment of school readiness of 5-year-old children in public pre-schools in 2012. In which, the development of children is measured by scores on five domains including physical health and well-being, social knowledge and competence, emotional health/maturity, language and cognitive development, and general knowledge and communication skills.

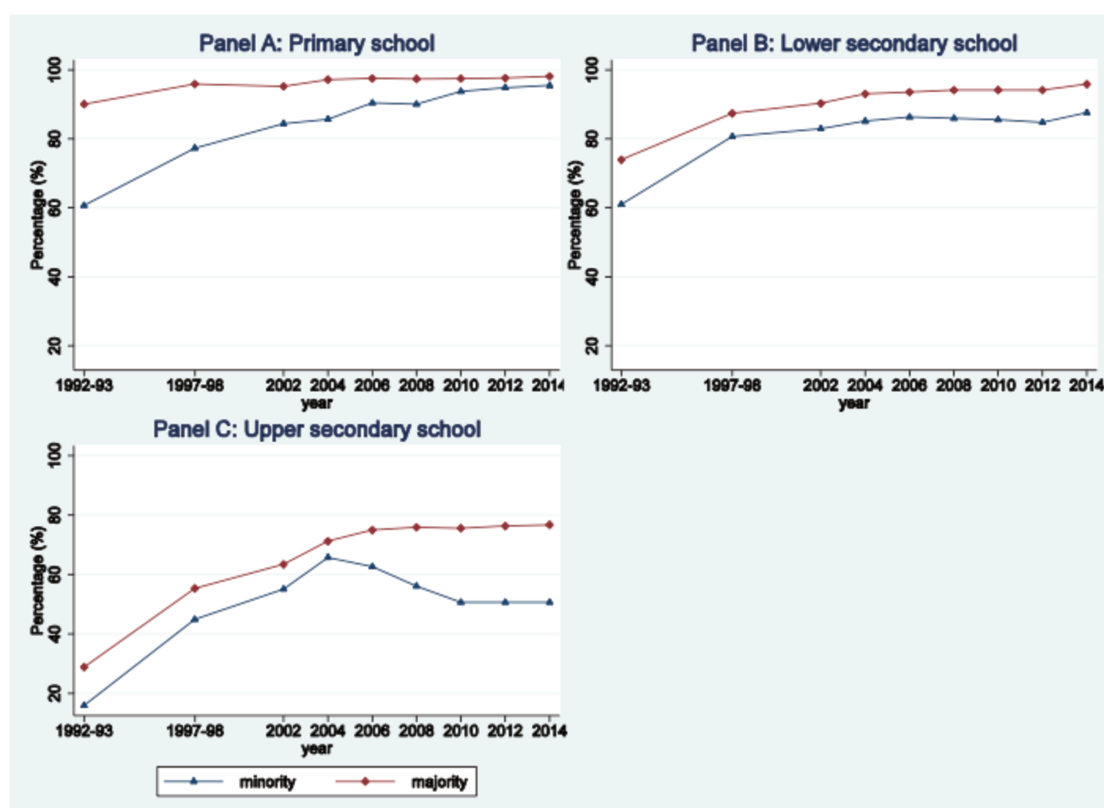
(Figure 0- 4). Girls had higher enrolment rate than boys and students in urban areas higher than in rural areas (Figure 0- 5). The low income, the minorities and the rural residents had lower educational levels completed than their counterparts from the high income, majority and urban residents (Kozel, 2014). Regarding learning achievement, girls, children with family's higher wealth index and higher parent's educational level had better performance (in Math and Vietnamese) than boys and children from family's lower wealth and lower parents' education (Figure 0- 6).

Figure 0- 3: Changes in net enrolment rate, by educational and income (percentage)



Source: extracted from World Bank (2014)

Figure 0- 4: Change in net enrolment rate, by education and ethnicity (percentage)



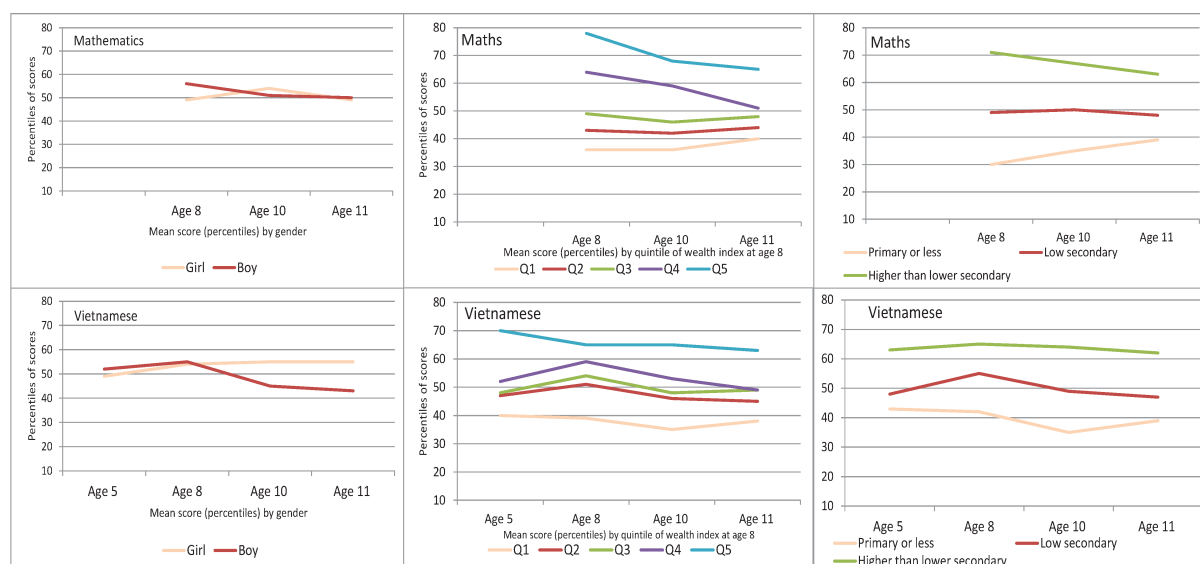
Source: extracted from Dang & Glewwe (forthcoming)

Figure 0- 5: Change in net enrolment rate, by education, gender and region (percentage)



Source: GSO (2016)

Figure 0- 6: Mean score (percentile) - in Maths and Vietnamese by gender, wealth and parents' education



Source: Rolleston et al. (2013)

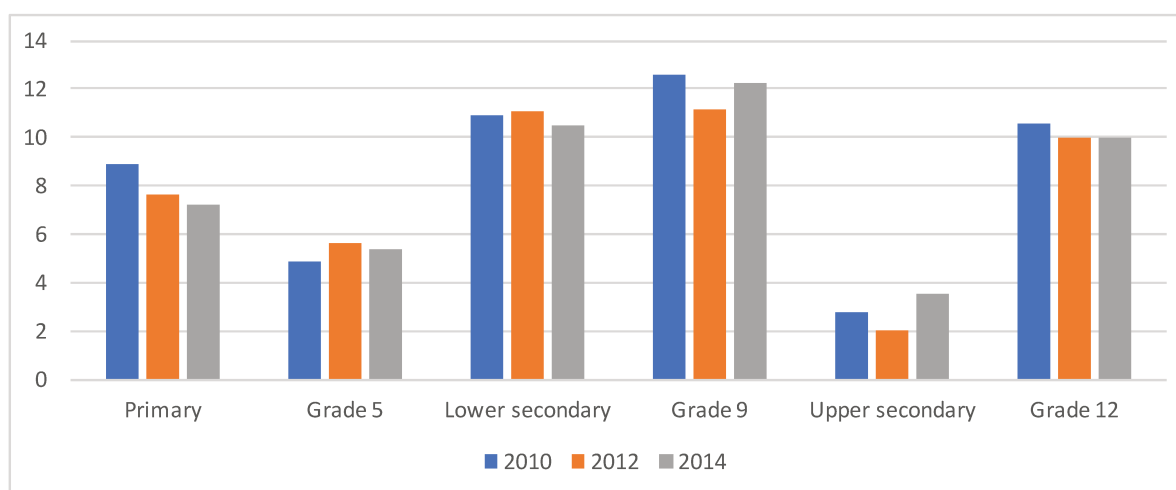
Another source of the gaps in skills is the country's trend in early school leaving. Vietnam was among ten countries⁸ in the world with more than 1 million out-of-school in 2005, and the number of students leaving school has risen sharply in recent years (UNESCO & Education for All, 2007). In 2009, 1.26 million children aged 5 to 14 either dropped out of school or had never attended (pre-primary/primary/secondary) school.⁹ Excluding the school leavers at the transition to upper secondary

⁸ The other countries were India, Nigeria, Pakistan, Burkina Faso, Côte d'Ivoire, Ethiopia, Kenya, Mali and Niger

⁹ Calculated based on the Population and Housing Census 2009

level (Grade 9), dropping out of lower secondary school in Vietnam accounts for the highest share of total dropouts in general education system and persists at about 11 percent (Figure 0- 7). As lower secondary education is a basic level of schooling that provides students with basic literacy and numeracy skills, the school leavers at this educational level would be short of these long-life skills for further learning and earning capacity.

Figure 0- 7: Dropout rate by educational level completed (percentage)



Source: Author's calculation based on VHLSS 2010, 2012, 2014

The gaps in skills cause a vicious circle of poverty. Children born into certain circumstances, for example, children in ethnic minorities or children with less educated parents or needy families, are more likely to rely on public schools to gain their education, thereby seeking for better well-being. In this regard, unequal opportunities for acquiring a good education and adequate skills may exacerbate inequalities confronted by the children and perpetuates their poverty.

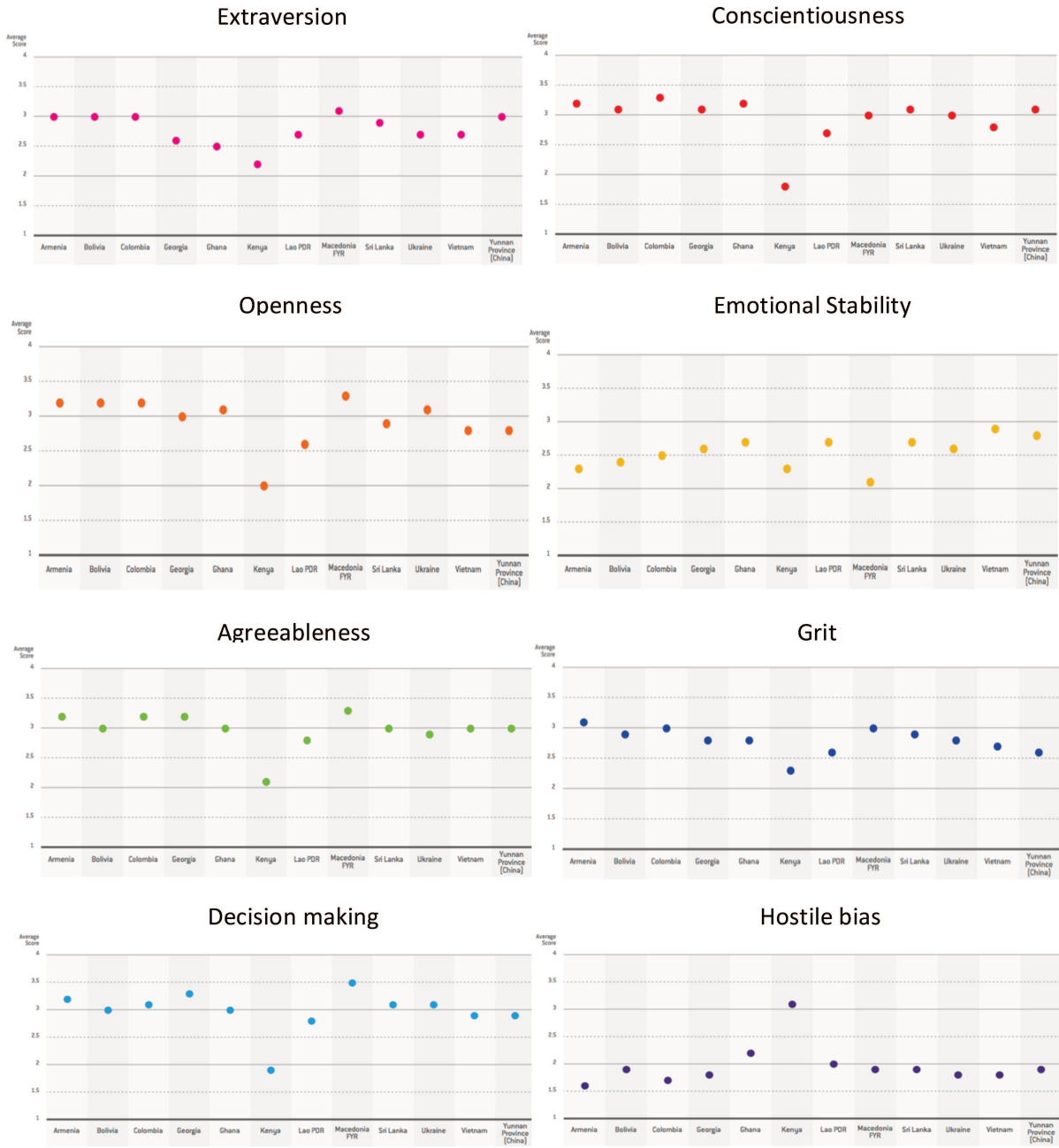
Second, one prominent weakness of the education sector is its failure in providing students with adequate non-cognitive skill. Non-cognitive skill is defined by personality traits or abilities that weakly correlate with measures of intelligence. In contrast to the impressive performance in cognitive skill, Vietnam's labour force score low in non-cognitive skills among developing countries. As shown in the results from the STEP¹⁰ household surveys by Roseth et al. (2016), Vietnam's scores¹¹ for Extraversion, Openness, Agreeableness, Decision Making and Hostile bias fall in the lowest one-third, and Conscientiousness and Grit in the lowest second and third. The trend excludes Emotional Stability which is highest for Vietnamese (Figure 0- 8). Analysis from the STEP employer survey also shows that

¹⁰ The World Bank's Skills Towards Employability and Productivity – Skill Measurement Program that consists of a household survey and an employer survey. The program collected information on supply and demand for skills in urban areas in 12 developing countries, including Bolivia, Colombia, Ghana, Lao PDR, Ukraine, Vietnam, and the Yunnan Province in China in 2012 and Armenia, Azerbaijan, Georgia, Kenya, and Macedonia FYR in 2013.

¹¹ Of which each is denoted by the second dot from the left side of each graph in Figure 0- 8.

most of the employers faced the lack of required skills and unavailability of skilled applicants when in hiring employees (Bodewig & Badiani-Magnusson, 2014).

Figure 0- 8: Non-cognitive achievement of Vietnam compared to other developing countries'



Source: Roseth et al. (2016)

The paradox

At present, despite significantly high labour productivity growth rate compared with other regional countries, Vietnam's labour productivity is among the lowest in the Asia-Pacific region.¹² The fact poses the question on what lies behind the paradox of very low productivity but very high level of cognitive skill and educational achievement. One possible explanation for the puzzle is the low level of non-cognitive skill. The labour force's skills constitute human capital which is, in turn, one primary source of labour productivity. As a result of world's rapid technological development and economic integration, the demand for high-skilled workers rises while demand for low-skilled workers declines globally. Vietnam is not the exception, particularly in the context of continuing the transition from an agricultural to a modern industrialised economy by 2020.¹³ While the labour force acquires high cognitive but low non-cognitive skill, the demand for highly non-cognitive-skilled has not been met. The non-cognitive skill thus has a higher return than the cognitive skill on labour market outcomes, causing the country's low productivity. Another reason for low productivity could be inequality and inequality of opportunity. Many pieces of evidence imply that beyond a small threshold of income inequality, a higher level of inequality negatively affects productivity (see, for example, Kim & Sakamoto, 2008; Pietro, 2014). In a more detailed view, the increase in inequality raises the decision power of the upper class in society. The higher the disparity, the more the upper class dominates in all life aspects. They are provided with opportunities to manipulate the system (the market, the media, the government and so on) to acquire things in their interests and to earn more income. Greater inequality and inequality of opportunity thus lead to a higher income of the upper class in an unproductive way, in other words, inequalities come with reduced productivity. Given Vietnam's rapidly ageing population is rapidly ageing and ongoing economic downturn,¹⁴ the low labour productivity is detrimental to the country's economic growth in coming years as well as the long-run prospect of Vietnam's development.

From the discussion as above, inequalities in education, the gaps in skills and insufficient non-cognitive skill apparently give rise to the concern on subsequent outcomes of human capital, economic growth and inequalities in Vietnam. It is not yet so clear about the sources of educational inequality and contrary outcomes on cognitive and non-cognitive achievement. These problems could result from one, some or all of more and more apparent school characteristics that negatively affect the quality education and school progress. The well-documented school fallouts, which may not be exhaustive, are the prevalence of private tutoring (Figure 0- 9), insufficiency of a teacher's official salary in covering his/her living expenditure; examination-oriented education system neglects other

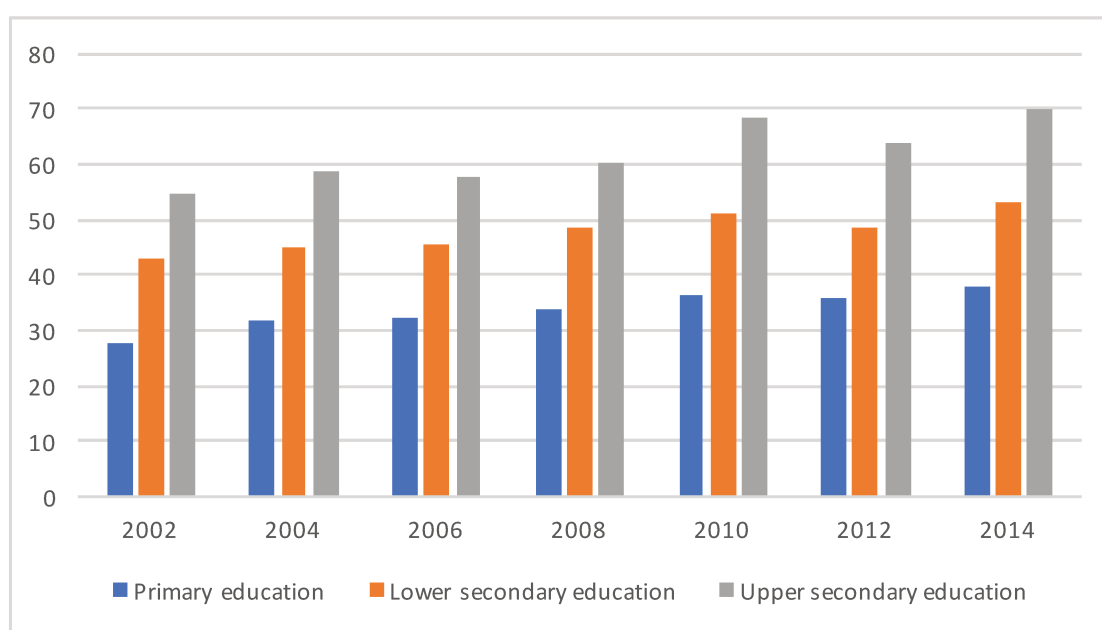
¹² According to the report by MPI-GSO (2016), the productivity in Singapore, Malaysia, Thailand, Philippines and Indonesia was nearly 18, 7, 3 and two times the level in Vietnam respectively in 2013.

¹³ articulated in the Vietnam's Socioeconomic Development Strategy for 2011–2020

¹⁴ Average GDP growth rate per annum decreased from 7.22 in the period 1993-2008 to 5.92 in 2009-2016.

aspects of pupils' skills; "achievement syndrome"¹⁵; corruption and abuse of power.¹⁶ The shortcomings do not assure the quality of learning and make the learning outcomes biased (Hoang, 2011). Each of these impediments can alone or in combination with others negatively affect the acquisition of skills and gaps in skills. For example, low compensation for teachers, private tutoring, corruption and abuse of power could make the inequality worse. Or the examination-oriented education system neglects the non-cognitive skill, either improving cognitive skill for children attending private tutoring or not improving; instead, the test scores are exaggerated as a consequence of corruption and abuse of power. Insufficient remuneration for teachers encourages teacher's private tutoring or teacher's corruption and abuse of power, thereby impact the quality of schooling and inequality.

Figure 0- 9: Attendance in private tutoring, by education (percentage)



Source: GSO (2016)

For these facts, there arises the need for significant reform of education sector, mainly the deal with the critical quality of general education schools, to provide equality of opportunities in education for children. There is the demand of policymakers on identifying the obstacles to acquiring adequate types and levels of skills. It is more pressing to study skill formation process for a thorough understanding of how innate and environmental factors (the family, school and community) interact in the early childhood to affect child's outcomes on skill and educational achievement and livelihood. Evidence on the roles of agents serves active intervention in the educational process for the equal

¹⁵ the tendency in which Vietnamese agencies, officials and teachers deliberately report exaggerated enrolment and graduation figures, regardless of learning outcomes (London, 2011).

¹⁶ Examples of which are the prevalence of exam cheating or extra classes in schools. These widen the disparities in learning outcome between poor and non-poor children.

opportunities and better adulthood well-being of children and the country's economic growth and equality of society as a whole.

2. Researches on skills and educational inequalities for Vietnam

The existing literature on skills formation, inequalities in education and well-being in Vietnam face three main limitations. First, most studies focused on establishing the correlation between social background and educational outcomes, or school environment and educational outcomes, but without addressing the interaction between social background and school environment, or the links between these three components. The reason for this limitation is that the existing studies rely either on household surveys (see, for instance, Glewwe, 2004a; Nguyen, 2006) that provide detailed information on the social background but very few information on the school environment. Or the studies use data from school surveys that measure cognitive skills and school environment but give inferior information on the social background as young children or teachers are the respondents (see, for example, World Bank, 2011). The second limitation is that they do not take into account the role of non-cognitive skills in the educational process (for example, Rolleston, 2014; Rolleston & Krutikova, 2014) and both cognitive and non-cognitive skills in the transmission of inequalities (e.g. Glewwe, 2012; Le & Baulch, 2012). The third limitation relates to the use of test scores as proxies for skills in the studies as mentioned above, which is associated with measurement error because test scores manifest not only skills but also other factors such as efforts or preferences (more details in the section on Definitions and measurement of skills below).

For these limitations, there has been no research on how skill formation determines outcomes on education and income or wage in Vietnam. Instead, present papers investigate a segment of the process. As regards the acquisition of skill, several studies imply that proper nutrition positively impacts the cognitive achievement, controlling for the characteristics of the child, child's household and community. In which, nutrition was measured by early childhood stunting (Le, 2011), undernutrition (Le & Behrman, 2017), growth between early and middle childhood (Georgiadis et al., 2016), and between middle childhood and early adolescence (Le & Tran, 2015). Some studies show that the gaps in skills are influenced by gender, ethnicity, parental education and wealth rather than extra classes (Le & Baulch, 2012). More particularly, ethnic gaps in cognitive skill were due to mother tongue, children's time allocation, years of schooling (Glewwe et al., 2012). Glewwe et al. (2012) also find the evidence of significant disparities even before primary education schooling (though the roles of these pre-school factors are not precisely defined). Meanwhile, the inequality of opportunities, inclusive of in education, was mainly due to the household head's education and expenditure (Nguyen & Vu, 2015; VASS, 2012). All these findings suggest that the inequality of opportunities leads to gaps in cognitive skill. However, these studies do not address the underlying mechanisms on inequality.

The studies on the link between skills and educational attainment and achievement come out together with the availability of the Young Lives in Vietnam (described in detail in the section on data below). This unique database provides rich data on children and their households over 15 years, notably including the information on children's cognitive and non-cognitive skills. Evidence from these studies reveals the positive association between cognitive skill and enrolment in school (see, for

instance, Rolleston, 2014; Rolleston & Krutikova, 2014) or between academic confidence¹⁷ as non-cognitive skill and learning progress (Rolleston et al., 2013). Moreover, the papers give the evidence of substantial differences in learning and learning progress, and that lower learning levels are associated with home background disadvantage and non-continuous patterns of enrolment.

No evidence of the link between skills and labour market results exists for Vietnam. Numerous studies use the household surveys to analyse the effect of education rather than the return of skills on employment, wage, earning or income. Examples are, for instance, Imbert, 2013; Liu, 2005, 2006; Moock, Patrinos, & Venkataraman, 2003; Oostendorp & Doan, 2013; Pham & Reilly, 2009; Sakellariou & Fang, 2014.. Neither the evidence on how skills and gaps in skills determine inequality in income or wage. Multiple empirical research tries to explain the overall inequality in income (Benjamin et al., 2017) or single dimension of inequality such as ethnic (Baulch et al., 2012; Pham & Reilly, 2009) or gender gap (Pham & Reilly, 2007). The decomposition analyses applied in the studies provide different findings on the explanatory power of the differences in individual/household endowment (inclusive of education) and differences in the return to the endowment on inequality.

The gap in the literature leaves it unclear why there is the paradox of high cognitive skill but low labour productivity for Vietnam. In my thesis, I aim to tackle the literature gap and overcome these limitations by using an original database collected by my research institution, the Centre for Analysis and Forecasting: Young Lives in Vietnam under Young Lives project. The Young Lives project is a longitudinal study of childhood poverty for understanding its causes and consequences to influence the formulation of public policies. It is led by a team in the Department of International Development at the University of Oxford. In association with other partners, the data are collected in four countries: Ethiopia, India (Andhra Pradesh), Peru and Vietnam. It follows over 15 years two cohorts of children, the Younger cohort and the Older cohort, born in 2001-02 and 1994-95 respectively. In this project, household surveys are conducted every four years and are combined with one school survey at age ten years.¹⁸ These data thus combine very rich information on the social background of the children, on their schooling experiences, on the school and teacher characteristics, as well as on their cognitive and non-cognitive, in a dynamic perspective. Taking advantage of these data, I will make following contributions on how skills interact with environmental factors to determine outcomes on education and livelihoods in Vietnam as follows. Firstly, my analysis of educational and labour outcomes takes into account both cognitive and non-cognitive skills. Secondly, I seek to investigate the underlying mechanisms of educational outcomes thanks to the combination of household and school surveys. Thirdly, I address the measurement errors due to the poor proxies for skills by letting skills dependent on test scores and other factors.

¹⁷ reflects students' feelings and perceptions about their academic competence one aspect of the academic self-concept (Rolleston et al., 2013)

¹⁸ see detailed information of these two datasets in the section "Data source" below

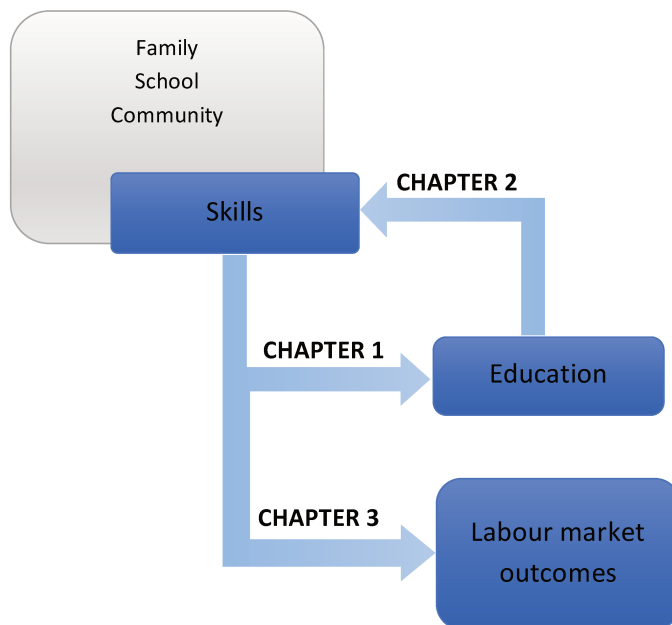
3. Objective, research questions and associated chapters

My thesis focuses on children in Vietnam and identifies what educational obstacles to levelling opportunities of children are. This research is expected to contribute its voice to the call for education reform aiming firstly at school quality for all. It seeks to examine the effects of school characteristics and social background on skill formation and transition to productive livelihood by addressing the three-dimension nexus between skills, education and labour outcomes (Figure 0- 10). In the nexus, I will tackle three sub-questions as below.

1) How do skills determine the dropping out for adolescents?

As discussed above, Vietnam has obtained outstanding performance in educational attainment and acquisition of skills. However, there have been some emerged problems of the education sector that raise the concern on the rising inequalities in well-being and declining economic growth shortly. Dropping out at lower secondary education is one among the problems because it leads to the lack of basic skills of the school leavers, who will face difficulties in seeking jobs and good earning. It therefore deserving of further investigating what lies behind children's decision to leave school. Moreover, there has been no evidence on how skills determine the school withdrawal. The phenomenon has not been studied in its nature of a dynamic process. Thus, a survival analysis of dropping out that takes into account skills will shed light on its role and other factors on dropping out over time.

Figure 0- 10: Thesis in a nutshell



2) How social background and school environment interact to determine the learning progress?

The children with social background disadvantages are more likely to face negative educational outcomes. In this case, intervention policies work if they can remove the obstacles to levelling opportunities in education. One of such policy is the country's implementation of full-day schooling

for primary education. While full-day schooling has been expanded massively since 2010, little evidence of its effect on inequality of opportunity is available. Therefore, the investigation of the relationship would serve the evidence-based policy of reforming the education sector. To this end, the research applies both descriptive analysis and value-added models using the data from Young Lives school survey in combination with household survey.

3) How do skills predict young people's labour market outcomes?

Following the evidence on how skills are formed to determine educational outcomes and inequality, this research aims to identify the relative importance of skills for the labour market outcomes. In addressing the question, it would elucidate the role of non-cognitive skill on livelihood. The paradox of high cognitive skill and low labour productivity as raised earlier can be explained, thereby giving a hint of the improvement of skills for higher labour productivity.

These research questions are addressed in three corresponding chapters. The summaries of the three chapters are provided in following sub-sections. Each review briefly presents the motivation of the research; the methodology applied, data and findings.

3.1. Chapter 1: Skills as the predictors of lower secondary dropping out in Vietnam: an estimation of Cox proportional hazard model

Lower secondary education for a long time ago has been considered as the educational level providing basic skills for further learning or developing the capability for future livelihoods (UNESCO¹⁹, 2014). For this reason, like other countries and by the commitment to the Education for All by 2015, primary and lower secondary education are compulsory in Vietnam. However only primary level has been nearly universal. Meanwhile, the enrolment in lower secondary level was about 90 percent in 2013. More concerned, the dropout rates at this educational level have remained at the rate of 11 percent since 2010. The dropouts, without basic skills, would face disadvantages in obtaining decent jobs and earnings, particularly in the context of increasing economic integration. Therefore, it has become an urgent need for policymakers to understand why adolescents drop out of lower secondary schools. In the international literature on factors of the dropping out, few studies consider the dynamic aspect of the phenomenon. Such evidence is even fewer for Vietnam, with limitations in the approaches and data employed. No study takes into account the role of skills in determining the phenomenon, mostly due to the unavailability of suitable data. Therefore, this chapter investigates the factors of dropping out at lower secondary level, using the Cox proportional hazard model and panel data for the Young Lives Older Cohort. It first identifies which type of skills are significant to dropping out, then examines how the predictive power of these skills for dropping out vary over time.

Though not addressing the causal effect of skills, the Cox models are justified for examining dropping in a dynamic context, dealing with the problem of censoring while allowing for time-variant

¹⁹ The United Nations Educational, Scientific and Cultural Organization

factors in the analysis. The chapter for the first time looks at the phenomenon in its nature of a dynamic process and gives the evidence on the predictive power of skills for dropping out over time.

The results of the analysis show that the cognitive skill, parents' education, adolescent's educational aspiration and time spent on household's chores are the most significant factors of dropping out at lower secondary education. The effects of Math, adolescent's educational aspiration and socioeconomic status (household's wealth index, household size and time children spent on household's chores) on the probability of dropping out at an earlier age magnify across later ages.

3.2. Chapter 2: Does full-day schooling reduces educational inequality in Vietnam?

Full-day schooling has been implemented in Vietnam to deal with the lack of instructional time in primary education regarding international standards. Lengthening instructional time is expected to improve learning achievement. Full-day schooling may impact educational inequality by filling the gap regarding instructional time between children with different family backgrounds. Indeed, well-off families use to compensate the lack of instructional time by private tutoring, which could give increase educational inequality. There is little evidence on the impact of full-day schooling on inequality of opportunity, which provides different findings. Therefore, this chapter investigates whether the massive development of full-day schooling in recent years can improve equality of opportunity in Vietnam by using the data from the Young Lives School Survey 2011-2012. Firstly, it examines the variation in instructional time and school resources between pupils with and without full-day schooling from different social backgrounds. It then investigates the learning progress associated with full-day schooling across social background by estimation of value-added models. To deal with the overrepresentation of the Young Lives children in the School Survey, these value-added models include the inverse Mills ratios. The ratios are obtained from the estimated Heckman two-step selection model of the probability for Young Lives children surveyed in 2002 to be included in the School Survey. Estimation of the models for subgroups of social backgrounds is also conducted to investigate the heterogeneity of the effect of full-day schooling.

Evidence implies that full-day schooling cannot narrow the inequality of opportunity in learning achievement at this stage of implementation. On the contrary, it contributes to rising the gap in learning progress among children of different social background.

3.3. Chapter 3: Does cognitive or non-cognitive skill matter in predicting labour market earnings: Evidence from Vietnam?

While there is consensus on the predictive power of both cognitive and non-cognitive skills in explaining earnings or wages, the relative importance of these skills is still debatable. Contrary to abundant studies that provide mixed evidence on the effects of these skills for developed countries, little evidence exists for developing countries. No such study is available for Vietnam. Besides, given the high level of cognitive skill but low non-cognitive skill in Vietnam, identifying which type of skills is

more significant provides important evidence to the policy for enhancing skills to improve labour outcomes. This chapter, therefore, aims to contribute to the controversial evidence, particularly for the developing world. It uses the Young Lives panel data for the Old cohort to examine the relationship between both types of skills and labour outcomes, including employment and earnings.

The chapter applies the structural equation modelling, which allows latent cognitive and non-cognitive skills, together with individual, family and community characteristics, determine the observed measures of skills and the person's choices on education and labour outcomes. The regression results imply the higher effect of non-cognitive than cognitive ability on employment and earnings. Of the sample, the low-skilled are more likely to be employed, with a higher predictive power of the cognitive skill. Among the employment, those with higher cognitive skill tend to work as self-employed, while the higher non-cognitively tend to be wage workers and often have higher earnings.

4. Skills: concepts and measurement

Given the substantial evidence of the predictive power of cognitive skills (rather than educational attainment) on individual earnings, distribution of income and economic growth, cognitive skills have become more and more important in promoting economic outcomes (Hanushek & Woessmann, 2008).

Cognitive skills, also called cognitive ability, cognition, intelligence or general mental ability. Psychologists define cognitive skills as *“ability to understand complex ideas, to adapt effectively to the environment, to learn from experience, to engage in various forms of reasoning, to overcome obstacles by taking thought”*, as quoted from one study in Almlund et al. (2011). Cognitive skills thus consist of multiple facets. Scores on different tests of cognitive skills such as IQ tests, achievement tests and grades, are highly correlated. Their correlation structure can be explained by a single factor labelled “g”. However, each test captures a distinct aspect of cognitive ability. IQ tests mainly measure the rate at which individuals learn (referred in psychology as ‘fluid intelligence’), while achievement tests primarily measure the acquired knowledge (‘crystallised intelligence’). The crystallised intelligence increases monotonically over most of the life cycle, while fluid reaches its peak at early adulthood then declines. Economists and psychologists interchangeably use IQ tests, achievement tests and grades to proxy cognitive ability.

Non-cognitive skills refer to the personal attributes not thought to be captured by the measures of cognitive skills. They also appear in other names in the literature such as personality traits, characters, socio-emotional skills, psychological competencies, soft skills. Each name emphasises some property of the attributes. The term “traits” implies the stability and inheritability of the attributes, while “skills” implies the attribute can be learned and altered. Both cognitive and non-cognitive skills evolve over the life cycle, and are malleable – in other words, can be changed over a period, but to different levels at different age and through different mechanisms. For this reason, this paper uses the terms “cognitive skills” and “non-cognitive skills” to reflect the malleability of these attributes.

There is no such “g” factor for non-cognitive skills as for cognitive skills. Measures of non-cognitive skills vary across studies. In many papers, non-cognitive skills are personality traits that weakly relate to measures of intelligence. The mostly used taxonomy of personality traits is the Five Factor Model (FF), or the Big Five in short, including Openness to experience, Conscientiousness, Extraversion, Agreeableness, Emotional stability. These factors are defined as follows (Almlund et al., 2011). *Openness to experience* is the tendency to be open to new aesthetic, cultural or intellectual experiences. *Conscientiousness* is the tendency to be organised, responsible and hardworking. *Extraversion* is an orientation of one’s interests and energies toward the outer world rather than being reserved. *Agreeableness* is the tendency to act cooperatively and unselfishly. *Neuroticism* is the tendency to experience emotional instability and psychological distress. (*Neuroticism* is the opposite of *Emotional Stability* - “predictability and consistency in emotional reactions, with absence of rapid mood changes”). Sometimes the Openness to Experience is replaced by *Autonomy*, which is the tendency to decide and degree of initiative and control. Each factor in the Big Five is a family of facets and related traits in the childhood and adulthood. For instance, the locus of control²⁰, self-esteem²¹ and self-efficacy²² are related traits that belong to Emotional Stability/Neuroticism, and Grit (persistence and passion for reaching long-term goals)²³, ambition and perseverance are the related traits that go with Conscientiousness. Some non-cognitive skills are used interchangeably due to the difficulty or inability to distinguish between these them. For instance, Self-esteem is also appeared as Self-concept in most cases due to the inability to (Byrne, 2002). Global Self-concept comprises academic and non-academic self-concepts. The latter consists of three facets which are Social, Emotional and Physical Self-concepts (ibid). Academic Self-concept relates to individual knowledge and perceptions about themselves in achievement situation (in the performance of all subjects or a subject) (Bong & Skaalvik, 2003). In many cases, Academic Self-concept can be considered as Self-efficacy because these constructs share many common features, but Bong & Skaalvik (2003) claim that “self-efficacy acts as an active precursor of self-concept development”. The Social self-concept includes Peer Relations, Parent Relations and relations with others.

The Big Five is argued to be unjustified because there is no scientifically grounded theory underpinning it. Besides, the Big Five has not developed to capture the heterogeneous nature of personality as well as their cultural variety across countries (Almlund et al., 2011). Alternative systems for the construct of non-cognitive skills include the related to the Big Five, and the extended, reduced or revised Big Five (as a result of linking the Big Five to measures of psychopathology) or a set of related traits.

The self-esteem and locus of control, collected by the large-sample longitudinal studies (e.g. the US National Longitudinal Surveys or the German Socio-Economic Panel) is the traditional alternative

²⁰ one’s belief about whether the determinants of one’s life events are largely internal or external

²¹ an individuals’ judgment of their own self-value or self-worth

²² one’s belief in their capabilities to achieve desired outcomes

²³ persistence and passion for reaching long-term goals

system that is related to the Big Five. These two traits and the self-efficacy, which play a fundamental role in human functioning both directly and indirectly (Bandura, 2006), have shown high validity.

Broadly defined, both cognitive and non-cognitive skills are not directly observed. They are inferred from the performance of tasks, for instance performing tests, using the analysis of latent variables and factors models. The performance of a test depends on a particular set of skills and other determinants (efforts, incentives), which firstly implies that a set of latent skills produce several scores for different tests. Secondly, test scores are imperfect measures of skills (because test scores also depend on the other determinants). As test scores depend on the same vector of skills, different test scores can be correlated. To ensure the validity of a set of test scores as measures of skills, a particular set of test scores captures a component of latent skills and not the others. Many test scores are taken to obtain the component to help controls for measurement error. The validation of measurement systems can be based on a priori grounds (which leads to the taxonomy of traits that became the Big Five described in detail below), or by any predictive validity regarding real-world outcomes (such as educational outcomes, earnings).

Measuring skills is challenging because it requires standardising for all other determinants of the observed outcome as mentioned above. Specifically, the estimates of skills are biased if not standardising for incentive, efforts and other skills. Also, there are false responses that arise from impression management and self-deception, the so-called “faking” that can bias the measures that proxy latent factors. Individuals tend to overstate their strengths and understate their weaknesses when responding to a self-reported questionnaire for recruitment. Or the examinees might try harder to gain higher achievement test score for promotion decision. Some evidence suggests that faking has a negligible effect on predicting job performance.

In the view that skills are inferred from task outputs, it is difficult to distinguish between cognitive and non-cognitive skills empirically. The IQ and achievement tests are also affected by non-cognitive skills such as motivation or efforts. Response to the self-reported questionnaire is influenced by the respondent’s strategy for responses that depend partly on their perception of gain from the reply and hence their cognitive skills. There are available econometric methods that enable isolating cognitive, non-cognitive skills and effect of environment and experience as well as accounting for measurement error.

5. Conceptualising skills in conventional economic models

In the last century, economists considered non-cognitive skills as the study focus of personality psychologists. Cognitive ability measured by achievement test scores are critical measures to assess individual's ability and predict their performance. However cognitive tests do not adequately capture the non-cognitive aspects of individuals, or non-cognitive skills, such as the personality traits, motivations, preferences, that are also significant in determining their outcomes on schooling, labour market and in other areas. Therefore, economists started to take into account the non-cognitive skills if they affect choices through preferences, constraints and information processing capabilities. The pioneer work in economics that analysed the impact of personality on earnings was Bowles & Gintis (1976). Other studies on the same (e.g. Mueser, 1979), and extended topic on the impacts of

personality on different domains, for instance on schooling or occupational choices as in Mueller and Plug (2006) follow this study.

Almlund et al. (2011) described in detail how to incorporate skills into standard economic models. According to which, the measures of non-cognitive skills can be defined as the performance outcome and effort that are inferred from solutions to a maximisation problem. The performance P on a task j depends on a vector of actions that rely on skills θ and effort e_j (where $\sum_{j=1}^J e_j = \bar{e}$) as follows:

$$P_j = v_j(\theta, e_j)$$

and is the solution to the maximisation problem subject to constraints on endowments and preferences.

In other words, the performance outcome and effort constitute the systems that produce measured non-cognitive skills in response to constraints, preferences and information. Skills are thus the endowments. The solutions to the maximisation problem generate a set of response functions or choices on action and effort subject to constraints.

Skills are not fixed. In a dynamic setting, skills, information, situation, actions evolve according to age, experience and investment. The skill formation technology by Cunha & Heckman (2007, 2009) defines the stock of skills in age $t + 1$ as:

$$\theta_{t+1} = (\theta_t, I_t, h_t), t \in T$$

in which: θ_t denotes individual's stock of skills θ at age t ; I_t denote the investments in period t include a broad range of actions such as parental care, nurturance, schooling, learning by doing and so on. Investments have dynamic effects, as skills evolve over times through investments. The skills possess the characteristic of self-productivity (skill begets skill) and cross-productivity (one skill begets another skill). Variable h denotes attributes of the situation in which the individual is placed.

In short, task performance is determined by a set of choices on actions to maximise agent's utility subject to constraints on endowments, inclusive of skills and effort, and preferences.

6. Data

6.1. Young Lives Household Surveys and School Survey in Vietnam

6.1.1. Young Lives Household Surveys

Young Lives in Vietnam follows 2000 children in Younger Cohort and 1,000 children in the Older Cohort. There have been five rounds of surveys conducted in 2002, 2006, 2009, 2013 and 2016. The most recent round (Round 5) has not been available up to now. The available four rounds of data correspond to the age of 1, 5, 8, 12, 15 of the Younger Cohort and the age of 8, 12, 15, 19, 21 of the Older Cohort.

Although the Young Lives household survey sampling is not designed to be representative of the Vietnamese population, it covers the diversity of children in the country in terms of a wide variety of attributes and experiences. These surveys use a sentinel-site sampling design comprising 20 purposely

selected sites chosen to represent diversity, but with a pro-poor bias (Nguyen, 2008). These sites are clustered within five provinces, which are Lao Cai, Hung Yen, Da Nang, Phu Yen and Ben Tre. Each province contains four sites, and each site is made up of one or two communes, totalling 36 communes in all. Sentinel site selection was made following a complex consultative process with a pro-poor selection rule based on each commune's poverty ranking. At the site level, children were selected randomly in 2001 such that the data are representative of the birth cohort at each site.

The Young Lives data provides rich information on diversified live aspects of children and their caregivers and families. Through the child, household and commune questionnaires, the project has collected comprehensive information on child characteristics, family, caregiver/parent characteristics and resources, their preferences and feelings, as well as school and community factors. Also, the distinctive feature of the database is the available information for measuring both child's cognitive and non-cognitive skill. The cognitive skill is measured by the results of the cognitive tests (e.g. Math, PPVT) administered to children. Definitions and measurement of all skills are provided in the section below. The availability of child's skills and rich information on child's surrounding environment (from family to community) over time enable the study of skill performance and how it relates to educational outcomes in a dynamic nature as in Chapter 1.

Also, the most recent round that is available – Round 4 (2013) for the Older Cohort children, who were at the age of 18-19, collects information on the cohort's labour participation (the last round was unavailable at the time of this research). In this Round 4, more comprehensive tests for non-cognitive skills were administered to the cohort. These tests were re-designed to adapt to a new situation where children were becoming adults and entering the labour market. These make available information on the key variables, which are labour outcomes (employment status and earnings), and associated cognitive and non-cognitive skills for investigating how skills affect labour outcomes in Chapter 3.

Among two cohorts, only data for the Older Cohort are relevant for the analyses in the first and third chapters. The first chapter examines determinants of dropping out of lower secondary school. This level of education involves children in the schooling age of 12 to 15. Thus, almost all of the dropouts are from the Older Cohort. The third chapter investigates the relative importance of skills in labour outcomes, which are expected to come out from the age of 15. In Round 4, the Younger Cohort was at the age of 12, thus had rarely participated in the labour market. Meanwhile, half of the Older cohort (at age 18-19) worked full-time. So, except for the analysis in Chapter 2 on the learning progress at the age of 10-11 that involves the Younger Cohort (explained in the next section on School Survey 2011-12), the Younger Cohort are excluded from the sample for analysis in the remaining chapters.

6.1.2. School Survey 2011-12

Vietnam School Survey is one among the surveys conducted in the four Young Lives countries in 2011-12 following the Young Lives Round 3 to capture the Younger Cohort children's experiences of schooling. It allows the link to longitudinal data from the household survey to the data on the schools attended by the Young Lives children in Vietnam. The combination of household, child and school-level data over time facilitates the policy-oriented studies on the relationship between child development process and their schooling experiences, including access, quality and progression.

Any class with at least one Grade-5 Young Lives Young Cohort child in the school located within the Young Lives sentinel sites was included in the School Survey sample. The sample was then expanded by randomly adding selected classmates of this (these) Young Lives child (children). In total, there are 20 pupils per class are included. The survey sample consists of 3284 Grade-5 pupils in 176 classes at 52 schools. Of which, 1125 are Young Lives children, and the rest are their classmates. The school's principal, class-teacher and pupils were all interviewed to provide a rich database on school, teachers and pupils. Besides the information on school and class facilities, pupils' tests on Mathematics and Vietnamese, pupils responded to questionnaire on the resources related to or devoted to their pursuit of schooling. At the same time pupils attended two tests on Mathematics and Vietnamese during the academic school year, they twice responded to attitude items designed for measuring non-cognitive skills. Meanwhile, teachers replied to attitude items for measuring teaching efficacy. Pupils' test scores and their non-cognitive scores, as well as teacher scores in two-time points in this survey, can be linked to the household survey to form the database for the analysis of learning achievement and progress mentioned in chapter 2 (the second research question).

6.2. Measures of cognitive and non-cognitive skills in the Young Lives Survey

Table 0- 1: Indices included in the composite index of cognitive and non-cognitive skill.

Age (Round)	No.	Cognitive skill	No of items	No.	Non-cognitive skill	No of items
12 (Round 2)	1	Pea Body Vocabulary Test	204	1	Self-efficacy/ Agency	5
	2	Math Test (60 minutes)	10	2	Self-esteem/ Pride	8
				3	Self-respect/ Inclusion	5
15 (Round 3)	1	Pea Body Vocabulary Test	204	1	Self-efficacy/ Agency	5
	2	CLOZE Test	24	2	Self-esteem/ Pride	6
	3	Math Test	29	3	Self-respect/ Inclusion	3
18 (Round 4)	1	Language Test (30 minutes)	24	1	General Self-efficacy	10
	2	Math Test (40 minutes)	28	2	General Self-esteem	8
				3	Parent Relations	8
				4	Peer Relations	8

The Young Lives in Vietnam give data to measure a variety of cognitive and non-cognitive²⁴ skills of children at different ages as listed in Table 0- 1 for the household surveys. The measures of cognitive skill provided in the School Survey 2011-12 are described in detail in Chapter 2. They include test

²⁴ Often referred to as psychological competencies in studies using Young Lives data

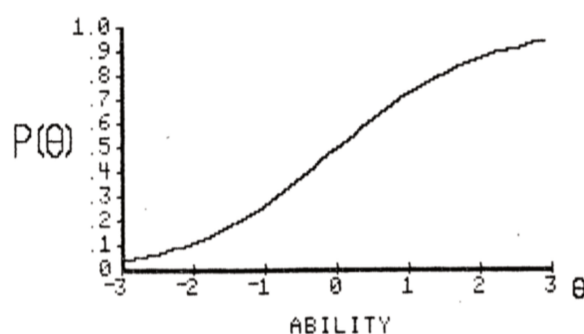
scores in Math and Vietnamese administered in the beginning and at the end of the school year. Each test consists of 30 multiple-choice items and lasted for 45 minutes.

The non-cognitive are measured by Self-efficacy/Agency, Self-esteem/Pride (from Round 2 to Round 4), Self-respect/Inclusion (Round 2 and Round 3), General self-efficacy, General self-esteem, Parent relations and Peer relations (Round 4 only). The definitions and computation of these skills are provided in the subsequent section.

6.2.1. Cognitive skills

The cognitive skills in the Young Lives surveys are measured by the test score in Numeracy/Mathematics (Math Test and CLOZE test), Peabody Picture Vocabulary Test (PPVT) and Language Test (Vietnamese). Specifically, two tests were administered to this cohort, including Math test and Peabody Picture Vocabulary Test in Round 1 to Round 3. CLOZE test was added in Round 3, and in Round 4, the PPVT was replaced by the Language test (in Vietnamese).

The single measures of the cognitive skill are standardized test scores in PPVT, Math, CLOZE and Language (Rasch scores) is computed using the Rasch model or the 1- parameter IRT model. Item Response Theory (IRT) is widely used in education, especially in the study of cognitive and non-cognitive skills. IRT mainly bases on the item characteristic curve (ICC), which describes the probability that a person succeeds on a given item (giving the correct answer to an individual test question). An individual responding to a test item has some amount of the latent ability. At each ability level, there is a probability that the individual provides a correct answer. Ability level varies across individual, with high ability individual having high probability and vice versa. The function of ability can be plotted as an S-shaped curve as below:



The likelihood of success on an item is a function of individual's latent trait or ability level and properties of the items, which is defined as follows:

$$P(\theta) = \frac{1}{1 + e^{-L}} = \frac{1}{1 + e^{-a(\theta - b)}}$$

where θ is the person's latent trait; the properties of the item include b - the difficulty parameter, a - discrimination parameter, and $L = a(\theta - b)$ is the logistic deviate. Discrimination parameter reflects the steepness of the item characteristic curve. It shows how well an item can distinguish between examinees having abilities below above the item location.

The Rasch model is the one-parameter IRT model. In this Rasch model, the discrimination parameter is fixed at a value of 1 for all items. It is specified as follows:

$$P(\theta) = \frac{1}{1 - e^{-1(\theta-b)}}$$

The Rasch model transforms raw scores onto a common scale for each test (namely Rasch scores), thus allowing tests to be comparable throughout the time.

6.2.2. Non-cognitive skills

6.2.2.1. Definitions

Self-efficacy/Agency

Self-efficacy has been widely studied theoretically and empirically since it was firstly introduced by Bandura in 1977.

“Perceived self-efficacy refers to beliefs in one’s capabilities to organize and execute the courses of action required to produce given attainments....Such beliefs influence the course of action people choose to pursue, how much effort they put forth in given endeavours, how long they will persevere in the face of obstacles and failures, their resilience to adversity, whether their thought patterns are self-hindering or self-aiding, how much stress and depression they experience in coping with taxing environmental demands, and the level of accomplishments they realize.” (Bandura, 1977)

Self-efficacy is defined as an individual’s belief in his/her capabilities to produce given attainments and to cope with adversity. It describes one’s confidence about whether his outcome is the result of his effort or the result of luck, fate or other’s intervention (Dercon & Krishnan, 2009). Items to measure self-efficacy were developed to capture different domains such as school, work and time use, for example ‘If I study hard at school I will be rewarded by a better job in future’. It is named as Agency by the Young Lives to distinguish this self-efficacy by Bandura from general self-efficacy scale in Young Lives Round 4.

General Self-efficacy: The construct of Self-efficacy in Round 4 applies the scale of Schwarzer & Jerusalem developed in 1995 to assess a general sense of perceived self-efficacy rather than specific behaviour. It has been adapted to many countries (including Hong Kong, India, Indonesia, Iran, Korea, Peru and Syria) and shown to be reliable and unidimensional across cultures. An example of a question from this scale is ‘I can always manage to solve difficult problems if I try hard enough’.

Pride/Self-esteem

Self-esteem (named as *Pride* in Round 3 to distinguish from new self-esteem scale administered in Young Lives Round 4) reflect how one values or like his or her self. Appearing in other names such as self-worth, self-regard, self-respect, self-acceptance, self-esteem is quantified as the sum of attitudes towards or evaluations of one’s self (Blascovich & Tomaka, 1991). It is distinct from self-efficacy in the

sense that self-efficacy is not perceived relative to others (Rosen & RTI International, 2010). Examples of the statements related to pride are 'I am ashamed/proud of my shoes' or 'I am proud of the work I have to do'.

Because Agency and Pride seemed to have performed well in previous rounds of Young Lives, the items related to agency and pride were included in Round 4 for comparable scales. Also, more comprehensive measures of general Self-efficacy and Self-esteem were also included in Round 4. The scale of General Self-esteem in Round 4 was administered using the Self-Description Questionnaire I. An example of a question measuring general self-esteem is 'In general, I like being the way I am'.

Self-respect and inclusion is another name for self-esteem (Blascovich and Tomaka 2011), which focuses on the social component of self-esteem (Dercon and Krishnan 2009). The statements used to measure self-respect revolve around the concepts of pride and sense of inclusion.

Parent relations is the indicator of how one gets along with their parents. An example of an item measuring positive relations with parents is 'I get along well with my parents'.

Peer relations indicates how one gets along with their peers. An example of an item measuring positive peer relations is 'I get along with other kids easily'.

6.2.2.2. Computation of the measures of non-cognitive skill

Measures of the non-cognitive skill are constructed based on respondent's degree of agreement or disagreement with some statements (see Appendix A for detailed statements and computation of each index). The degree of agreement is based on a 4-point scale ranging from "strongly disagree" to "strongly agree". For the computation of each single measure, all the items are recoded to reflect positive statements. The degree of agreement with these statements is also recoded to range from strong disagreement to strong agreement. Each single measure of the non-cognitive skill is the average of normalised responses for all the statements used to construct the skill. All measures of the non-cognitive skill are reliable because the Cronbach's alpha²⁵ for the non-cognitive indexes we obtain from our analysis is at least 0.7, except for Agency/Self-efficacy and Inclusion (see Table 0- 2 for the reliability of the indexes). However, these indices are still taken into consideration in studies using the same source of data for 4 countries inclusive of Vietnam, such as Dercon and Krishnan (2009), Dercon and Sanchez (2011), Dercon and Sanchez (2013) and Sanchez (2013). Most recently Dercon and Sanchez (2013) "obtain values of 0.89 and 0.95 for the self-efficacy and self-esteem indicators" in the pooled sample of all Young Lives countries. Therefore, Agency/Self-efficacy and Inclusion are kept for panel observation across round, and as additional measures for new self-efficacy and general self-esteem in Round 4.

²⁵ Commonly used measure of scale reliability. Specifically, the Cronbach's alpha, which is defined as 'the square of correlation between the measured scale and the underlying factor', would indicate whether different items in the scale are measuring the same latent variable. The critical value for the reliability of the scale is 0.7.

Table 0- 2: Cronbach's alpha for the non-cognitive scales

	2006 (Round 1)	2009 (Round 3)	2013 (Round 4)
Agency/Self-efficacy	0.3984	0.3221	0.4291
Pride/Self-esteem	0.6818	0.7033	0.6955
Inclusion	0.5119	0.4746	
General self-efficacy			0.7085
General self-esteem			0.7012
Parent relations			0.8107
Peer relations			0.7462

References

- Almlund, M., Duckworth, A. L., Heckman, J., & Kautz, T. (2011). Personality Psychology and Economics. In *Handbook of the Economics of Education* (Vol. 4, pp. 1–181). Elsevier. Retrieved from <http://linkinghub.elsevier.com/retrieve/pii/B9780444534446000018>
- Bandura, A. (1977). *Social Learning Theory*. Oxford, England: Prentice-Hall.
- Bandura, A. (1993). Perceived Self-Efficacy in Cognitive Development and Functioning. *Educational Psychologist*, 28(2), 117–148.
- Bandura, A. (2006). Guide for constructing self-efficacy scales. In *Self-efficacy beliefs of adolescents* (F. Pajares & T. Urdan, Vol. 5, pp. 307–337). Information Age Publishing. Retrieved from <https://www.uky.edu/~eushe2/Bandura/BanduraGuide2006.pdf>
- Baulch, B., Pham, H. T., & Reilly, B. (2012). Decomposing the Ethnic Gap in Rural Vietnam, 1993–2004. *Oxford Development Studies*, 40(1), 87–117. <https://doi.org/10.1080/13600818.2011.646441>
- Benjamin, D., Brandt, L., & McCaig, B. (2017). Growth with equity: income inequality in Vietnam, 2002–14. *The Journal of Economic Inequality*, 15(1), 25–46. <https://doi.org/10.1007/s10888-016-9341-7>
- Blascovich, J., & Tomaka, J. (1991). Measures of Self-Esteem. In *Measure of Personality and Social Psychological Attitudes* (John P. Robinson, Phillip R. Shaver, Lawrence S. Wrightsman, Vol. 1 of Measures of Social Psychological Attitude). Gulf Professional Publishing.
- Bodewig, C., & Badiani-Magnusson, R. (2014). *Skilling up Vietnam: Preparing the Workforce for a Modern Market Economy*. The World Bank. <https://doi.org/10.1596/978-1-4648-0231-7>
- Bong, M., & Skaalvik, E. M. (2003). Academic Self-Concept and Self-Efficacy: How Different Are They Really? *Educational Psychology Review*, 15(1), 1–40. <https://doi.org/10.1023/A:1021302408382>
- Byrne, B. M. (2002). Validating the measurement and structure of self-concept: Snapshots of past, present, and future research. *American Psychologist*, 57(11), 897–909. <https://doi.org/10.1037/0003-066X.57.11.897>
- Cunha, F., & Heckman, J. (2007). The Technology of Skill Formation. *American Economic Review*, 97(2), 31–47. <https://doi.org/10.1257/aer.97.2.31>
- Cunha, F., & Heckman, J. (2009). *The Economics and Psychology of Inequality and Human Development* (No. w14695). Cambridge, MA: National Bureau of Economic Research. Retrieved from <http://www.nber.org/papers/w14695.pdf>
- Dang, H.-A., & Glewwe, P. (forthcoming). Well Begun, But Aiming Higher: A Review of Vietnam's Education Trends in the Past 20 Years and Emerging Challenges. *Journal of Development Studies*.
- Dercon, S., & Krishnan, P. (2009). Poverty and the Psychosocial Competencies of Children: Evidence from the Young Lives Sample in Four Developing Countries." *Children, Youth and Environments*, 19(2), 138–163.
- Georgiadis, A., Benny, L., Crookston, B. T., Duc, L. T., Hermida, P., Mani, S., ... Behrman, J. R. (2016). Growth trajectories from conception through middle childhood and cognitive achievement at age 8 years: Evidence from four low- and middle-income countries. *SSM - Population Health*, 2, 43–54. <https://doi.org/10.1016/j.ssmph.2016.01.003>

- Glewwe, P. (2004a). An Investigation of the Determinants of School Progress and Academic Achievement in Vietnam. In *Economic growth, poverty, and household welfare in Vietnam* (Paul Glewwe, Nisha Agrawal, and David Dollar). Washington, DC: World Bank.
- Glewwe, P. (2004b). An Overview of Economic Growth and Household Welfare in Vietnam in the 1990s. In *Economic growth, poverty, and household welfare in Vietnam* (Paul Glewwe, Nisha Agrawal, and David Dollar). Washington, DC: World Bank.
- Glewwe, P., Qihui, C., & Katare, B. (2012). What Determines Learning Among Kinh and Ethnic Minority Students in Vietnam? An Analysis of the Round 2 Young Lives Data. Young Lives Working Paper 80. Retrieved from <http://younglives.qeh.ox.ac.uk/publications/WP/determines-learning-kinh-ethnic-minority-students-vietnam/wp80-what-determines-learning-among-kinh-and-ethnic-minority-students-in-vietnam>
- GSO, G. S. O. (2016). *Result of the Vietnam Household Living Standards Survey 2014*. Hanoi, Vietnam: Statistical Publishing House.
- Hanushek, E. A., & Woessmann, L. (2008). The Role of Cognitive Skills in Economic Development. *Journal of Economic Literature*, 46(3), 607–668. <https://doi.org/10.1257/jel.46.3.607>
- Hoang, T. (2011). Khủng hoảng giáo dục: Nguyên nhân và lối ra trước thách thức toàn cầu hoá [Education crisis: reasons and solutions in the context with globalization challenge]. In T. Hoang (Ed.), *Giáo dục: Xin cho tôi nói thẳng*.
- Holsinger, D. B., & Jacob, W. J. (2008). Education Inequality and Academic Achievement. In D. B. Holsinger & D. B. Jacob, W. James (Eds.), *Inequality in education: comparative and international perspectives*. Dordrecht: Springer.
- Imbert, C. (2013). Decomposing the Labor Market Earnings Inequality: The Public and Private Sectors in Vietnam, 1993-2006. *The World Bank Economic Review*, 27(1), 55–79. <https://doi.org/10.1093/wber/lhs009>
- Kim, C., & Sakamoto, A. (2008). Does Inequality Increase Productivity? Evidence From U.S. Manufacturing Industries, 1979 to 1996. *Work and Occupations*, 35(1), 85–114. <https://doi.org/10.1177/0730888407311975>
- Kozel, V. (2014). *Well Begun but Not Yet Done: Progress and Emerging Challenges for Poverty Reduction in Vietnam*. Washington, DC.
- Le, T. D. (2011). Height and Cognitive Achievement of Vietnamese Children. *World Development*, 39(12), 2211–2220. <https://doi.org/10.1016/j.worlddev.2011.04.013>
- Le, T. D., & Baulch, B. (2012). Do Extra Classes Improve Cognitive Test Scores? Evidence from Vietnam. Young Lives Working Paper 87.
- Le, T. D., & Behrman, J. R. (2017). Heterogeneity in predictive power of early childhood nutritional indicators for mid-childhood outcomes: Evidence from Vietnam. *Economics & Human Biology*, 26, 86–95. <https://doi.org/10.1016/j.ehb.2017.02.002>
- Le, T. D., & Tran, N. T. M. T. (2015). Growth in Middle Childhood and Early Adolescence, and its Association with Cognitive and Non-cognitive Skills at the Age of 15 Years: Evidence from Vietnam. Young Lives Working Paper 138.
- Liu, A. Y. C. (2005). Impact of Economic Reform on Wage Structure in Vietnam. *East Asian Economic Perspective*, 16(2), 46–57.

- Liu, A. Y. C. (2006). Changing wage structure and education in Vietnam, 1992-1998: The roles of demand. *The Economics of Transition*, 14(4), 681–706. <https://doi.org/10.1111/j.1468-0351.2006.00266.x>
- Moock, P. R., Patrinos, H. A., & Venkataraman, M. (2003). Education and earnings in a transition economy: the case of Vietnam. *Economics of Education Review*, 22(5), 503–510. [https://doi.org/10.1016/S0272-7757\(03\)00039-6](https://doi.org/10.1016/S0272-7757(03)00039-6)
- MPI-GSO, M. of P. and I.-G. S. O. (2016). *Năng suất lao động của Việt Nam: Thực trạng và giải pháp [Productivity in Vietnam: situation and solution]*.
- Nguyen, N. N. (2004). Trend in education sector. In P. Glewwe, N. Agrawal, & D. Dollar (Eds.), *Economic growth, poverty, and household welfare in Vietnam*. Washington, DC: World Bank.
- Nguyen, P. L. (2006). Effects of social class and school conditions on educational enrollment and achievement of boys and girls in rural Viet Nam. *International Journal of Educational Research*, 45(3), 153–175. <https://doi.org/10.1016/j.ijer.2006.11.008>
- Nguyen, P. N. (2008). An assessment of the Young Lives Sampling Approach in Vietnam, Young Lives Technical Note n°4.
- Nguyen, T. K. T., & Vu, T. T. T. (2015). Reducing Inequality of Human Opportunities in Vietnam: A Challenge for Social Progress. Paper under the Mekong Economic Research Network funded by International Development Research Centre (IDRC), Canada (project 105220). Retrieved from <http://mernetwork.org/wp-content/uploads/2015/08/SG-4-Vietnam-Kim-Thai-final.pdf>
- Oostendorp, R. H., & Doan, Q. H. (2013). Have the returns to education really increased in Vietnam? Wage versus employment effect. *Journal of Comparative Economics*, 41(3), 923–938. <https://doi.org/10.1016/j.jce.2012.12.002>
- Oxfam. (2017). *Even it Up: how to tackle inequality in Vietnam*. Labor - Social Publishing House. Retrieved from https://www.oxfam.org/sites/www.oxfam.org/files/file_attachments/bp-vietnam-inequality-120117-en.pdf
- Pham, T.-H., & Reilly, B. (2007). The gender pay gap in Vietnam, 1993–2002: A quantile regression approach. *Journal of Asian Economics*, 18(5), 775–808. <https://doi.org/10.1016/j.asieco.2007.04.001>
- Pham, T.-H., & Reilly, B. (2009). Ethnic wage inequality in Vietnam. *International Journal of Manpower*, 30(3), 192–219. <https://doi.org/10.1108/01437720910956727>
- Pietro, W. R. D. P. (2014). Productivity Growth and Income Inequality. *Journal of Economics and Development Studies*, 2(3). <https://doi.org/10.15640/jeds.v2n3a1>
- Rew, R. W. (2008). Provincial, Ethnic, and Gender Disparities in Education: A Descriptive Study of Vietnam. In D. B. Holsinger & W. J. Jacob (Eds.), *Inequality in education: comparative and international perspectives*. Dordrecht: Springer.
- Rolleston, C. (2014). Learning profiles and the ‘skills gap’ in four developing countries: a comparative analysis of schooling and skills development. *Oxford Review of Education*, 40(1), 132–150. <https://doi.org/10.1080/03054985.2013.873528>
- Rolleston, C., James, Z., Pasquier-Doumer, L., & Tran, N. M. T. (2013). Making Progress: Report of the Young Lives School Survey in Vietnam. Young Lives Working Paper 100. Retrieved from http://www.younglives.org.uk/publications/WP/making-progress-report-of-the-young-lives-school-survey-in-vietnam/yl-wp100_rolleston

- Rolleston, C., & Krutikova, S. (2014). Equalising opportunity? School quality and home disadvantage in Vietnam. *Oxford Review of Education*, 40(1), 112–131. <https://doi.org/10.1080/03054985.2013.875261>
- Rosenberg, M. (1965). Society and the adolescent self-image. *Princeton University Press*.
- Roseth, V. V., Valerio, A., & Gutiérrez, M. (2016). *Education, skills, and labor market outcomes: results from large-scale adult skills surveys in urban areas in 12 countries* (Main report). Washington, D.C.: World Bank Group. Retrieved from <http://documents.worldbank.org/curated/en/397871467995623739/Main-report>
- Rotter, J. B. (1966). Generalized expectancies for internal versus external control of reinforcement. *Psychological Monographs: General and Applied*, 80(1).
- Sakellariou, C., & Fang, Z. (2014). The Vietnam reforms, change in wage inequality and the role of the minimum wage. *Economics of Transition*, 22(2), 313–340. <https://doi.org/10.1111/ecot.12037>
- Schwarzer, R., & Jerusalem, M. (1995). Optimistic Self-Beliefs as a Resource Factor in Coping with Stress. In S. E. Hobfoll & M. W. Vries (Eds.), *Extreme Stress and Communities: Impact and Intervention* (pp. 159–177). Dordrecht: Springer Netherlands. Retrieved from http://link.springer.com/10.1007/978-94-015-8486-9_7
- UNESCO (Ed.). (2014). *Teaching and learning: achieving quality for all*. Paris: UNESCO.
- UNESCO, & Education for All (Eds.). (2007). *Education for All by 2015: will we make it?* Paris: UNESCO.
- VASS. (2012). Opportunities for Children in Vietnam. (Vietnam Academy of Social Sciences) Background paper for the 2012 Poverty Assessment.
- World Bank. (2011). *Viet Nam: High Quality Education for All by 2020 (Vol. 2): Analytical report* (Working paper No. 68092).
- World Bank. (2014). *Taking Stock, July 2014: An Update on Vietnam's Recent Economic Developments*. Hanoi, Vietnam. Retrieved from <https://openknowledge.worldbank.org/handle/10986/19987>

Appendix

A. Variables included in scales of non-cognitive skill

Index (Based on [Reference])	Round 2 (Age of 12)	Round 3 (Age of 15)	Round 4 (Age 18)
Self-esteem/ Pride (Rosenberg Self- esteem Scale (Rosenberg, 1965))	<ol style="list-style-type: none"> 1. I feel proud to show my friends or other visitors where I live 2. I am ashamed of my clothes (recoded to positive) 3. I am ashamed of my shoes (recoded to positive) 4. I feel proud of the job my mum/dad does 5. I am often embarrassed because I do not have the right books, pencils and other equipment for school (recoded to positive) 6. I am worried that I don't have the correct uniform (recoded to positive) 7. I am proud of my achievements at school 8. I am embarrassed by/ashamed of the work I have to do (recoded to positive) 9. The job I do makes me feel proud 	<ol style="list-style-type: none"> 1. I am proud of my shoes or of having shoes 2. I feel my clothing is right for all occasions 3. I am proud of my clothes 4. I am often proud because I do have the right books, pencils and other equipment for school 5. I am proud that I have the correct uniform 6. I am proud of the work I have to do 	<ol style="list-style-type: none"> 1. I am proud of my clothes. 2. I feel my clothing is right for all occasions. 3. I am proud of my shoes or of having shoes. 4. I am proud of the work I have to do.
General self- esteem			<ol style="list-style-type: none"> 1. I do lots of important things 2. In general, I like being the way I am

(Self-description Questionnaire I)			<p>3. Overall, I have a lot to be proud of</p> <p>4. I can do things as well as most people</p> <p>5. Other people think I am a good person</p> <p>6. A lot of things about me are good</p> <p>7. I'm as good as most other people</p> <p>8. When I do something, I do it well</p>
Self-Efficacy/Agency (Rotter, 1966; Bandura, 1993)	<p>1. If I try hard, I can improve my situation in life</p> <p>2. Other people in my family make all the decisions about how I spend my time (recoded to positive)</p> <p>3. I like to make plans for my future studies and work</p> <p>4. If I study hard at school I will be rewarded by a better job in future</p> <p>5. I have no choice about the work I do - I must work (recoded to positive)</p>	<p>1. If I try hard I can improve my situation in life</p> <p>2. Other people in my family make all the decisions about how I spend my time (recoded to positive)</p> <p>3. I like to make plans for my future studies and work</p> <p>4. If I study hard at school I will be rewarded by a better job in the future</p> <p>5. I have choice about the work I do – I must do this sort of work</p>	<p>1. I like to make plans for my future studies and work.</p> <p>2. Other people in my family make all the decisions about how I spend my time.</p> <p>3. If I try hard, I can improve my situation in life.</p> <p>4. If I study hard at school I will be rewarded by a better job in future.</p> <p>5. I have no choice about the work I do - I must do this sort of work.</p>
General Self-Efficacy (Schwarzer & Jerusalem, 1995)			<p>1. I can always manage to solve difficult problems if I try hard enough</p> <p>2. If someone opposes me, I can find the means and ways to get what I want</p> <p>3. It is easy for me to stick to my aims and accomplish my goals</p>

			<p>4. I am confident that I could deal efficiently with unexpected events</p> <p>5. Thanks to my resourcefulness, I know how to handle unforeseen situations</p> <p>6. I can solve most problems if I invest the necessary effort</p> <p>7. I can remain calm when facing difficulties because I can rely on my coping abilities</p> <p>8. When I am confronted with a problem, I can usually find several solutions</p> <p>9. If I am in trouble, I can usually think of a solution</p> <p>10. I can usually handle whatever comes my way</p>
Inclusion	<p>B. When I am at shops/market I am usually treated by others with fairness and with respect</p> <p>C. Adults in my STREET/VILLAGE treat me worse than other children my age (recoded to positive)</p> <p>D. The other children in my class treat me with respect</p> <p>E. Other pupils in my class tease</p>	<p>1. The other children in my class treat me with respect</p> <p>2. Pupils in my class never tease me at school</p> <p>3. Adults in my community treat me as well as they treat other children at my age</p>	

	me at school (recoded to positive) F. My teachers treat me worse than other children (recoded to positive)		
Peer relations (Self-description Questionnaire I)			<ol style="list-style-type: none"> 1. I have lots of friends 2. I make friends easily 3. Other kids want me to be their friend 4. I have more friends than most other kids 5. I get along with other kids easily 6. I am easy to like 7. I am popular with kids of my own age 8. Most other kids like me
Parent relations (Self-description Questionnaire II)			<ol style="list-style-type: none"> 1. I like my parents 2. My parents like me 3. My parents and I spend a lot of time together 4. I get along well with my parents 5. My parents understand me 6. If I have children of my own, I want to bring 7. My parents are easy to talk to 9. My parents and I have a lot of fun together

Chapter 1: Skills as the predictors of lower secondary dropping out in Vietnam: An estimation of Cox proportional hazard model

Abstract

Lower secondary education is critical in providing essential skills for further learning and developing the capability for future livelihoods. Though compulsory in Vietnam, the enrolment rate was 90 percent in the school year 2013-2014. The dropout rates at this educational level have remained at the rate of 11 percent since 2010, causing more concerned on the disadvantages faced by the school leavers in obtaining decent jobs and earnings. Given few studies on the dynamic aspect, and little evidence, with limitations in the approaches and data employed, of dropping for Vietnam, this chapter investigates factors of the decision to drop out at lower secondary school, using the Cox proportional hazard model and panel data for the Young Lives Older Cohort. The chapter for the first time looks at the phenomenon in its nature of a dynamic process and gives the evidence on the predictive power of skills for dropping out over time. The results of the analysis show that the cognitive skill, parents' education, adolescent's educational aspiration and time spent on household's chores are the most significant factors of dropping out at this educational level. The effects of Math, adolescent's educational aspiration and socioeconomic status (household's wealth index, household size and time children spent on household's chores) on the probability of dropping out at an earlier age magnify across later ages.

1. Introduction

Reducing dropping out of secondary education has become a policy concern nowadays. Primary schooling helps people to escape from illiteracy and innumeracy which are forms of human insecurity and barrier to human development (Sen, 2003). Secondary education prepares adolescents for appropriate learning and life skills learning. Of which the lower secondary schooling provides the most efficient way to obtain fundamental skills. These skills, including basic literacy and numeracy skills, serve as the foundation for further learning to develop the capacity for decent jobs and lives (UNESCO²⁶, 2014). Because of progressive decline in primary school age out-of-school rates (to only nine percent worldwide in 2014 according to UIS & GEMR²⁷ (2016)) and growing demand for qualified workers in today's increasingly globalised world, dropping out at the lower secondary level has taken priority over primary in education research and policymaking. Compulsory schooling has been extended to include the lower secondary (in four to seven-year duration, inclusive of the upper secondary in some countries) in 80% of all countries in 2009 (UIS, 2011) and then nearly worldwide in 2014. However, the global rates of out-of-school children aged between 12 and 15 have increased since 2010, with every one of six (16 percent) not attending schools in 2014 and thus ineligible for entering upper secondary or vocational education (UIS & GEMR, 2016).

Vietnam provides an interesting case study of early school leaving, especially for the lower secondary education in a developing country. Basic schooling in Vietnam includes primary and lower secondary education. As a member of the Education for All, Vietnam aimed at universal primary and lower secondary education by 2015 (National Education for All Action Plan 2003-2015). Only the target for the former has been achieved, with the net enrolment rate of 98 percent, the completion rate of 92 percent in 2012 and nearly all graduates continued to lower secondary level. Meanwhile, the corresponding net enrolment and completion rates for the latter were 88 and 83 percent. The number of adolescents aged 10 to 14 who were lower secondary dropouts predominated in general education at the rate of 11.8 percent in 2009 (MOET, 2013). The dropout rates almost remained at about 11 percent in 2010, 2012 and 2014 (own calculations from the Vietnam Household Living Standard Surveys in 2010, 2012 and 2014). The school leavers would not achieve lower secondary qualification, which could put them in the situation of disadvantages compared with the graduates in obtaining jobs and earnings. This trend gives rise to the policy concern on possible inequality in well-being as well as the lack of skilled human resource for the target of Vietnam to turn the economy into a 'modern industrialised' one by 2020 (as articulated in Vietnam's Socioeconomic Development Strategy for 2011–2020). Understanding the reasons why adolescents leave schools at lower secondary level is therefore crucial to policy-making in the current context.

International literature on the factors that determine the decision to drop out is profuse. Researchers have applied many approaches to analyse school drop-out at the level of secondary education (see De Witte et al., 2013 for a review). Empirical research has identified diversified

²⁶ The United Nations Educational, Scientific and Cultural Organization

²⁷ UNESCO Institute for Statistics and Global Education Monitoring Report

predictors of dropping out of school, including the characteristics of the students and their families, and the contextual factors related to their schools and communities. Cognitive and non-cognitive skills are among the critical student-related factors of dropping out. Intuitively, the skills acquired at the lower secondary are in fact necessary for students to pass the exam or be eligible for enrolling in the upper secondary education. The evidence on the predictive power of skills on school leaving is mixed. The majority of evidence shows that poor academic achievement that manifests a low level of cognitive skills strongly predicts school leaving. Meanwhile, the remaining implies that age or other pull factors including high school pregnancy or employment fully explain the correlation between academic ability and dropping out (Allensworth, 2005; Rumberger, 2004). A shred of evidence is available on equivocal effects of non-cognitive skills, of which self-esteem²⁸ mostly appears (Rumberger, 1987). Other non-cognitive skills ever considered as negative predictors of school withdrawal include low self-efficacy²⁹ (Alivernini & Lucidi, 2011), internalised locus of control³⁰ (Coneus, Gernandt, & Saam, 2011), academic self-concept³¹ (Coroner, Linnanmäki, & Aunio, 2014).

Despite numerous dropout studies, few take into consideration the dynamic nature of dropping out. Most of the available studies examine whether or not students leave school at a given time, which can be addressed using cross-section data with historical information on individual's enrolment. The static approach fails to deal with the bias due to the problem of censoring that arises because some students drop out just before or after the time of the survey. The strategy also ignores the variation in time of the factors (for instance, the evolvement of skills over time) as well as in the effects of the factors that can be meaningful for the analysis. The phenomenon is thus not fully explored, and the drawing inferences do not provide reliable evidence for timely policy intervention and planning.

Contrary to the abundant international literature, the empirical literature for Vietnam is fewer with limitations in the methodologies and data applied for the analysis of school drop-out. First, none of the existing studies on school leaving or completion for Vietnam investigates the impact of skills, especially non-cognitive ones, on the probability of dropping out. This research gap mainly, but not necessarily, results from the unavailable data at the times of the studies. The existing papers either exclude skills (Baulch, Vu, & Nguyen, 2012) or use poor proxies of skills. Examples of such proxies are grade completed (Bui, 2011; Vo & Trinh, 2005), non-standardized test scores (Glewwe, 2004; Nguyen, 2006), and caregiver's assessment of child's class performance (Le & Tran, 2013). Neither studies look at any non-cognitive skill as determinants of dropping out. Thus, there is no proof as to what extent skills are predictive of dropping out that could be useful hints for restricting the phenomenon in

²⁸ one's subjective estimation of his/her own worth

²⁹ one's belief in their capabilities to achieve desired outcomes

³⁰ one's belief that his/her life outcomes are the result of his/her own actions. This is contrary to the externalized locus of control which attributes outcomes to fate/luck

³¹ mental representation of one's competencies in academic domains

Vietnam. Second, the analyses of the determinants of dropout have not been done longitudinally, therefore not yet taken into account its dynamic nature.

Therefore, this chapter inspects the determinants of dropping out at lower secondary education. It applies survival analysis using the panel data for the children born in 1994-1995 from the Young Lives in Vietnam, a 15-year study on child poverty in developing countries. Specifically, longitudinal factors of dropping out, inclusive of child's cognitive and non-cognitive skills, are analysed employing the Cox proportional hazard models. Although the models cannot identify the causal effects of the factors, the method is the most appropriate for investigating determinants of dropouts in a dynamic context. In comparison with traditional logit or probit models, the Cox model examines when students drop out, dealing with the problem of censoring and allowing for time-varying factors in the analysis. Identifying when students drop out is also informative of the time for intensified investments in children for better adult outcomes. Bearing in mind the pros and cons as mentioned above, the paper aims at policy implication on restraining dropping out of lower secondary school and brings the following value-added to previous studies of the same research questions. Firstly, it provides evidence on the predictive power of skills on dropping out for the first time in Vietnam. Secondly, it first examines the determinants of dropping out in Vietnam in its nature of a dynamic process. Estimation results show that the cognitive skill, parents' education, adolescent's educational aspiration and time spent on household's chores are the most significant factors of dropping out of lower secondary school. The effects of Math, adolescent's educational aspiration and socioeconomic status (household's wealth index, household size and time children spent on household's chores) on the hazards at an earlier age magnify across later ages.

The structure of the paper is as follows. Section 2 provides state of the art in the academic literature on school dropout, with a focus on theoretical and empirical analytical framework for developing countries. Section 3 presents an overview of education in Vietnam, including issues and problems that the system has faced in recent years. Section 4 describes the data and estimation strategy. Section 5 provides the quantitative analysis of the determinants of school drop-out. Section 6 concludes and suggests policy recommendations.

2. Literature review

2.1. International studies on the determinants of school dropout

Theoretically, whether or not an individual continues schooling is the decision on investment in education. According to the framework of Becker and Tomes (Becker & Tomes, 1979, 1986), the investment in a child's education is determined by the maximisation of parents' utility that takes into account their income, consumption and the expected return to the child's education. This theoretical model predicts that in a context of financial constraint, the investment is determined by the child's endowments (i.e. cultural and genetic attributes), parental and public investments (and willingness in financing the investments) in the child's development, as well as unsystematic components (market luck). These components (child's endowments, parental and public investments) are respectively referred to as the factors related to the child (or the student), the student's family, and student's

school and community in the empirical studies on the determinants of early school leaving. Or the factors of dropping out are grouped into the “push factors” and “pull factors” by some researchers depending on whether the factors “push” students towards or “pull” them away from schools.

Empirically, child endowments include gender, race or ethnicity and cognitive and non-cognitive skills and educational aspiration. Regarding gender, it is relatively evident that male is more likely to leave school than female (Rumberger, 2004). Alternative findings show that the gender effect can be insignificant (Cataldi, Laird, & KewalRamani., 2009; Kaufman, Alt, & Chapman, 2004). Similar disputable proof exists for ethnicity and race. For example, higher propensity to drop out is observed for black students (Bradley & Renzulli, 2011) or coloured students (Fleisch et al., 2012).

The proxy for cognitive skill is academic achievement. The academic performance is measured either by test scores (in such subjects as math, numeracy, literacy, language, grade point average), or by success in school and exit exams, or grade retention. Most studies indicate that students with poor academic achievement in high school are more likely to drop out (Battin-Pearson et al., 2000; Branson, Hofmeyr, & Lam, 2014; Janosz, LeBlanc, Boulerice, & Tremblay, 1997; Korhonen et al., 2014; Rumberger, 1987). But some studies found no effect, inclusive of Allensworth (2005). This paper is the only one that addresses causal relationship between grade retention (or repetition) and dropping out based on a matched-samples approach (see more details on this article on the review of dropout studies using survival analysis below). Allensworth (2005) suggests that repetition (based on standardised testing) is a stronger predictor, and academic achievement could indirectly affect school withdrawal through retention: Students with higher cognitive skills have a lower chance of retention, hence would be less likely to drop out. The effect of retention on dropping out is significant after a cut-off age, as such “pull” factors as teenage pregnancy or employment (Allensworth, 2005; Rumberger, 2004). In some contexts, the effect of academic achievement become less significant when controlling for other factors such as socioeconomic status and school quality in the context of South Africa (Branson et al., 2014), and learning difficulties and school stress in Finland (Korhonen et al., 2014),.

Proxies for the non-cognitive skill could be self-esteem, self-efficacy, the locus of control or academic self-concept. Existing evidence provides univocal skill effects on school leaving is much fewer for the non-cognitive than for the cognitive skills. *Self-esteem* is the most common non-cognitive skill that is examined in the literature. This factor appeared in one of the earliest frameworks of school dropout - the frustration-self-esteem model in Finn (1989), which explain dropout as a developmental process. The model presumes that school’s failures in providing sufficient instruction and emotional environment cause unsuccessful school outcomes. The poor school performance causes frustration or embarrassment, which harms student’s self-esteem, thereby leads to problematic behaviour, inclusive of leaving school. Empirical evidence shows that self-esteem is a significant factor of secondary dropping out: the dropouts have lower self-esteem (Kaplan, Peck, & Kaplan, 1995; Rumberger, 1987). Such effect of self-esteem on the likelihood to drop out is significant but negligible in No et al., (2016) in the context of Cambodia. *Self-efficacy* is another proxy for the non-cognitive skill that determines to drop. It is clear from the research findings that self-efficacy affects motivation and achievement in children and adolescents (Schunk & Pajares, 2001). Following the social cognitive theory reviewed in Schunk & Zimmerman (2014), self-efficacy, which is “people’s

beliefs in his/her ability to learn or perform actions at designed level”, can influence the choice of activities, effort, persistence, resilience and learning. Thus, students with high self-efficacy are more likely to survive longer when encountering academic difficulties. Examples of empirical evidence of the impact of self-efficacy on school leaving are Caprara et al., (2008) and Alivernini & Lucidi (2011) which use longitudinal designs (structural equation models) in the context of Italy. Results from Caprara et al., (2008) reveal that self-efficacy in high school contributed to both high school grades and school dropout, controlling for prior academic achievement. The lower the decline in self-efficacy, the higher the grades and likelihood of persisting for high school students, controlling for family background. In their study, Alivernini & Lucidi (2011) find that self-efficacy influenced on the intention to drop out of high school through its effect on motivation and academic achievement. Other non-cognitive skills considered as predictors of school withdrawal include locus of control³² (Coneus et al., 2011) and academic self-concept³³ (Korhonen et al., 2014). Specifically, the students with externalised locus of control (believe that their outcomes result from their actions), and high academic self-concept (high self-perception of their academic ability) are more likely to remain in school.

Child’s educational aspiration is also a decisive factor of school leaving (Rumberger, 2011). The student with a higher level of educational aspiration faces a lower risk of dropping out, controlling for exogenous factors such as geographic residence family background and ability (Rumberger, 1983). No studies have identified a causal effect of educational aspiration or underlying mechanism of the effect.

The proxies for parents’ investments in their child’s development are mostly family background or socioeconomic status. The most common measures of the family background are parents’ income, or education or parents’ aspiration for child’s education. Of these factors, parental income and education are the most common measures that strongly correlate with the likelihood of early school leaving. While researchers most often agree on the protective role of parental education against their children’s withdrawal from schools, the association between family or parental income and the likelihood to drop out is most debatable. Evidence suggests the influence of income on dropout (e.g. Yi et al., 2012; Sabates et al., 2013), with varying definitions of income. The predictive power of parental income is either nullified or valid for particular threshold or specific population groups (De Witte et al., 2013; review in Branson et al., 2014). A shred of evidence is suggestive of the influence of parents’ (or caregiver’s) educational aspiration on their children’s educational aspiration in determining the dropout risk. More generally, parental level of education and expectation for children impact the probability to leave school might be twofold, including the effect on children’s educational aspiration and the effect on children’s abilities subject to the level of their care and involvement in their children’s schooling. In addition to parental education, aspiration and income, other measures of family background such as parents’ employment or occupation are sometimes also predictive of early school leaving (Marks et al., 1999).

³² One’s belief that his/her life outcomes are the result of fate/luck (external) or of his/her own actions (internal).

³³ A pupil’s self-perception of his or her own academic performance, which is at least in part a relative concept, defined in relation to a peer group (Liu & Wang, 2008; Marsh & Hau, 2003).

The family structure represented either by the number of household members or siblings, single or two-parent household is another proxy for parental investments. Students from larger families or with single parenthood are more likely to leave schools early. The significance of these factors might depend on the level or quality of parental involvement and care.

The proxies for community's investments in the child development are residential (geographical) location or region. The influence of peers in the community plays a significant role in the student's decision to drop out. It was not the peer effect that draws the attention of scholars previously. Instead, the friends' educational aspiration and expectations are more appealing. For a student living in a community where there are high-aspiring or high-achieving friends in the student's environment, or most friends remaining in schools, his/her likelihood to drop out is smaller (Rumberger, 1983; Cooper et al., 2005). However to what extent and in what ways the student's friends and peers impact dropping out is unclear (Rumberger, 1987). Similarly, the community with a lot of employment opportunities might be counted as a factor pulling students out of school (Marks et al., 1999), though the magnitude of the factor relies on how the employment match student's circumstances.

The decision to leave school is the last stage of a dynamic, cumulative and multidimensional process. The complexity of the phenomenon is demonstrated by a variety of reasons provided by dropouts and the variation in the importance of the reasons over time. It is hard, or even impossible to infer causality from all factors on school leaving, except for family background and innate skills which are predetermined before the decision. Moreover, some factors are impacted by other factors, causing the inexact magnitude of the factors. In most cases, empirical evidence suggests rather than proves a causal relationship between the factor and dropping out (Rumberger, 2004; Smeyers, 2007).

Survival analysis versus other approaches

Researchers have conducted dropout studies employing either cross-sectional or panel data. The latter has been recommended (Willett & Judith, 1991) and increasingly used (De Witte et al., 2013) because it deals with the shortcomings of the former. The framework of Becker and Tomes shows that the decision to leave school the last phase of a dynamic process that begins early. Meanwhile, cross-section data only provides information to estimate to what extent explanatory variables predict dropping out at a particular point in time. It neither takes into consideration the variation in time of the coefficients nor measures the changes in the magnitude of the factors over time, thus only allows for examining school leaving statically. For this reason, the inspection of withdrawal out should rely on panel data to enable the investigation of dropping out as a dynamic, cumulative and multidimensional process.

A majority of dropout studies investigate whether students leave schools. In other words, the analyses in these studies involve splitting the sample into those who have dropped out and have not at a particular cut-off period. Such a dichotomous approach has three main limitations. First, the method ignores the problem of censoring, which results from a part of the sample not dropping out during the time of the survey, but (just) before or after the cut-off period. The exclusion of the already dropouts in the dropouts and the inclusion of the going-to-be dropouts in the non-dropouts gives rise to the bias in the analysis of the phenomenon. Second, the dichotomous approach treats all dropouts as a group, ignoring when students dropped out. The time of dropping out, which is informative for

the periods of time in student's schooling demanding more concentration or more investments regarding sources and efforts for better adult outcomes. In these regards, event history analysis which estimates hazard probability or survival analysis is a more straightforward but sound alternative statistical method to the dichotomous one (Willett & Judith, 1991). It is a longitudinal analytic technique that permits researchers to assess the transition from being students to dropouts while including all variables that record the students' different situation over time. By revealing valuable information on the occurrence of events, the survival analysis enables the inspection of both whether students drop out and when students are more likely to drop out, accounting for the censoring. Third, to take into account time-varying factors or to allow the factors to vary over time, the researchers applying dichotomous approach use/construct the factors' values comparable to a point in time or the averaged values or rate of change in values across time. The conclusions from the approach should, therefore, be attached to the time frame of the surveys involved in the analysis. Meanwhile, the survival analysis provides a unique method for including both static and time-varying factors, as well as varying effects of the factors.

A Cox proportional hazard model is only descriptive and does not address the causal identification of dropping out. It comes with the endogeneity due to the relationship between these explanatory variables and dropping out (for instance, the propensity to withdraw from school causes student's poor performance or low self-esteem) as well as the measurement errors in measuring skills. There are also the difficulties in determining the magnitude of factors due to the multicollinearity between explanatory variables (for example, academic performance is affected by family background). Although not addressing all the problems of the static model, particularly the causality identification, survival analysis is a "powerful" method that allows for analysing dropouts as a dynamic process. It provides the tool to address the research questions on whether and when students drop out as well as which factors predict dropping out, treating the problem of censoring. Dropouts in survival analysis are "not only of dropout (in all of its guises) but also of graduation, stop-out, and other transition at the high-school, college, and post-graduate levels" (Willett & Judith, 1991).

In fact, survival analysis is widely applied in economics but on other topics than in educational outcomes (see review in DesJardins et al., 2006). Few studies on high school dropouts have employed survival analysis for developed countries (Allensworth, 2005; DesJardins et al., 2006; Kalmijn & Kraaykamp, 2003; Vizcain, 2005) and developing countries (No et al., 2016; Paura & Arhipova, 2014; Woldehanna & Hagos, 2015). Kalmijn & Kraaykamp (2003) examine determinants of dropout and downward mobility (i.e. shifting from the general to vocational track, a lower level of the secondary school system) for secondary students in the Netherlands. They conduct an event-history model with multinomial logistic regressions employing longitudinal data. Estimates demonstrate a significant effect of being in the majority/minority of the population. Immigrants' children have a higher probability of dropping out than Dutch's, which results from the gaps in individual ability and parental resources. For any failures experienced at school, immigrants' children tend to leave school while Dutch's tend to move downward. Allensworth (2005) examines the effect of retention resulting from high-stakes testing (named the grade promotion) on dropout based on a matched-samples approach. He compares the students in Chicago before and after implementation of the eighth-grade promotional gate. The study combines many methods (hierarchical generalised linear model

combined with instrumental variables, logit, probit, survival analysis), of which survival analysis is used to explore whether the policy impacts the timing of dropping out at the age of 15 to 17. Findings indicate that retention due to the grade promotion raised dropout rate but to a smaller degree than the retention traditionally initiated by teachers, and this adverse effect of the policy was unrelated to the timing of leaving school. Vizcain (2005) inspects which factors associated with high school withdrawal using survival analysis and logistic regression. Data for the study comes from high school male students from Hispanic/Latino over a five-year span in the state of Florida. The findings suggest that age, retention, courses available, academic performance and behaviour problem (disciplinary suspension) are significant factors of dropping out. About 18 years old was the most hazardous time for dropping out. DesJardins et al. (2006) study the effects of stop-out, dropout, re-enrollment, on graduation from college. The authors apply multiple spells-competing risks hazard model using longitudinal data on different groups of students (who first time enrolled in University of Minnesota-Twin Cities campus in 1984, 1986 and 1991). Results show the significant effects of an initial stop-out on subsequent stop-outs, of the length of initial spell on the likelihood of following events. Also, the impact of race on the differences in student performance becomes insignificant when accounted for income, entry age and high school performance.

All three studies for developing countries (No et al., 2016; Paura & Arhipova, 2014; Woldehanna & Hagos, 2015) apply proportional hazard model in analysing the causes of school dropout. Paura & Arhipova (2014) identify the predictors of the first-year students' dropout (the survival time being the study duration (in months) in the first year). The study relies on real data of engineering program in the Latvia University of Agriculture. It finds that secondary school score, gender and low motivation to study the program are primary reasons for dropping out. Employing the data from two rounds of the Young Lives for Ethiopia and hazard model, Woldehanna & Hagos (2015) inspect the impact of economic shocks, which result from drought, crop or livestock failure, or death or illness of household members, on dropout rates for primary students in Ethiopia. Results indicate that the shocks in the previous round of the survey, late schooling age negatively impact the probability of completing primary education. No et al. (2016) looks at both primary and high school level in a Cambodian province over three years. The study shows that divorced parents, relationships with friends and late school entry significantly increased the odds of dropouts for students in grade 1-4. For students in grades 4-9, dropping out was significantly associated with grade repetition and academic achievement.

None of the dropout studies using survival analysis as mentioned above does not examine the magnitudes of factors across time.

2.2. Studies on dropping out for Vietnam

The studies on dropping out in Vietnam has been conducted using dichotomous approach (logit/probit models). The approach, as discussed earlier, has not addressed the problem of censoring as well as the dynamic aspect of the phenomenon. Regarding the framework of Becker and Tomes, the dropout studies for Vietnam have not sufficiently taken into consideration child's skills, one crucial component of child's endowment.

Most of these studies have employed the nationally representative data from the Vietnam (Household) Living Standard Surveys (Nguyen, 2006; Vo & Trinh, 2005; Bui, 2011). The studies share two common features. First, their findings are representative. Second, these analyses classify a child as having dropped out if he/she has not entirely enrolled in schools in the 12 months before the survey, given he/she used to enter in school sometime before. Third, they have included a rich set of factors that are proxies for child's endowment, family's investment and community's investment. Vo & Trinh (2005) and Bui (2011) use data from different waves of VLSSs have provided the evidence on effects of other common factors on the probability of dropping out of school. These factors include, as proxies for child's endowment, age, school level, time spent on working. For family background, the proxies are parental education, household expenditure, the cost of schooling, family size, the number of siblings. Proxies for community's investment are the region, urban/rural location and public spending on education. Compared to Vo & Trinh (2005) and Bui (2011), the value-added of Nguyen (2006) involves the inclusion of test scores (as the additional proxy for children's skills) and school quality (as the other proxies for community investment) in the analysis for more reliable estimation results. The author observed that children from poor households in better-off villages (without Government's financial support as in disadvantaged areas) faced the high risk, and girls with average learning progress possibly had less chance. It is noted that there is possible measurement error in using the test scores (ranging from 0 to 10) in the analysis of Nguyen (2006). These test scores are not standardised across schools (to control for measurement errors for comparable scores), thus do not provide consistent or reliable measures of skills.

Remaining studies on the determinants of school dropouts in Vietnam use data from the Young Lives, a longitudinal study of childhood poverty in four countries (Baulch et al., 2012; Le & Tran, 2013). Both studies focus on children who have dropped out before entering the upper secondary level. Dropping out was identified based on the information on child's enrolment at age 14-15 (by Round 3 of the Young Lives Surveys). The studies apply logistic models that incorporate contemporaneous child-related, family-related and community-related variables and lagged variables on the family's socioeconomic status. Baulch et al. (2012) in turn consider two separate groups of children who have dropped out of school by age 14-15, the time of the Young Lives Round 3 (in 2009). The first group include the children that have completed lower secondary completion (and already enrolled in upper secondary schools), and the second are the children that had not completed lower secondary schools. They find that for the first sub-sample, the strongest predictors were parental education and the amount of paid or unpaid work undertaken by the child in the previous round in 2006. For the second sub-sample, however, only mother's education level, together with ethnicity are important factors behind the likelihood of drop-out before lower secondary completion. While Baulch et al. (2012) do not take into account child's skills, Le & Tran (2013) incorporate this child's endowment using parents' assessment of their children's class performance. Le & Tran (2013) quantitatively and qualitatively analyse early school leaving in general and leaving right after lower secondary education. The authors employ the same data on Young Lives children as in Baulch et al. (2012), additionally counting their siblings. Logistic regression results indicate poor class performance has the highest predictive power for the propensity to drop out. Then comes the poor socioeconomic status – poverty and parental education less than lower secondary level. Findings from the qualitative analysis reaffirm the principal

effect of class performance while attributing the perception of the value of schooling rather than poverty to the dropping out, especially for ethnic children.

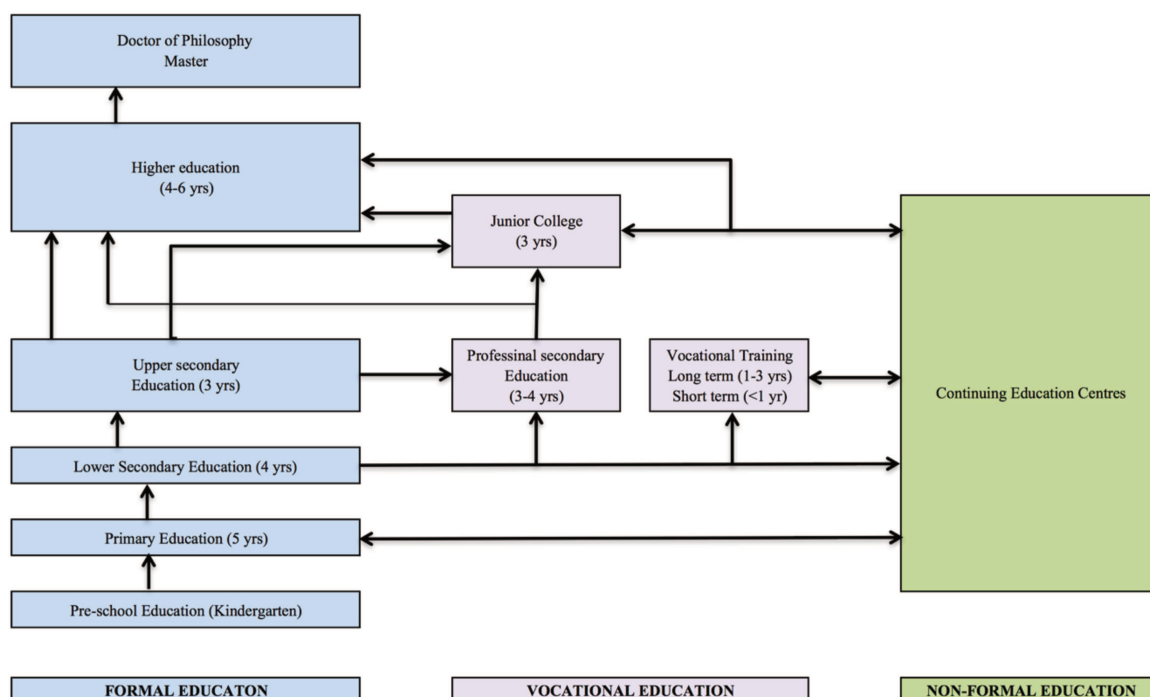
So, the dropout studies for Vietnam have not entirely taken into consideration child's skills, one among the most critical components of child's endowment. Only poor proxies of skills such as highest grade completed, or parents' assessment of their children's class performance, or non-standardized test scores present in the papers. No studies take into account student's non-cognitive skill. Except for Le & Tran (2013), all of these studies emphasise the significance of family background, for instance, parental education, rather than skills in predicting school leaving. The dichotomous method applied in the studies ignores the problem of censoring and time dependent factors and changing effects of the factors over time.

The current paper, therefore, applies survival analysis, specifically Cox proportional hazards models, to investigate the determinants of school leaving in a dynamic process, taking into account both cognitive and non-cognitive skills using the panel data from the Young Lives in Vietnam. The longitudinal and multidimensional approach would provide a comprehensive picture of dropout in its nature and suggest policy implications for limiting the dropout rate in Vietnam.

3. Education in Vietnam: an overview

3.1. Education system

Figure 1- 1: Vietnam's education system



Source: Adapted from the website of the Vietnam Ministry of Education and Training (www.moet.gov.vn).

Three intertwined blocks (formal, vocational and non-formal education) in Figure 1- 1 describe the education system in Vietnam. Primary school is compulsory and consists of Grades 1 to 5. Children start Grade 1 at the age of 6.

It typically takes four years to complete lower secondary education (Grades 6–9). Having finished Grade 9, young people have to pass an entrance exam to get into upper secondary schools (Grades 10–12). Students who had completed lower secondary education but failed or did not take the upper secondary school entrance exam are denied further schooling in the regular academic track under formal education unless they try again and are successful in the exam the following year. In 2014, the dropout rate at the transition to upper secondary school was 10.5 percent (GSO & UNICEF, 2014). The dropouts of the lower secondary schools and dropouts at the transition to upper secondary level can go into secondary technical and vocational education and training, which can consist of short or regular programs (see Figure 1- 1). Others follow non-formal education, partly made up of Continuing Education Centres that provide opportunities for those who are not able to take full-time and regular courses. It is not common for young people with qualifications from the technical and vocational training schools and the Continuing Education Centres to go onto universities.

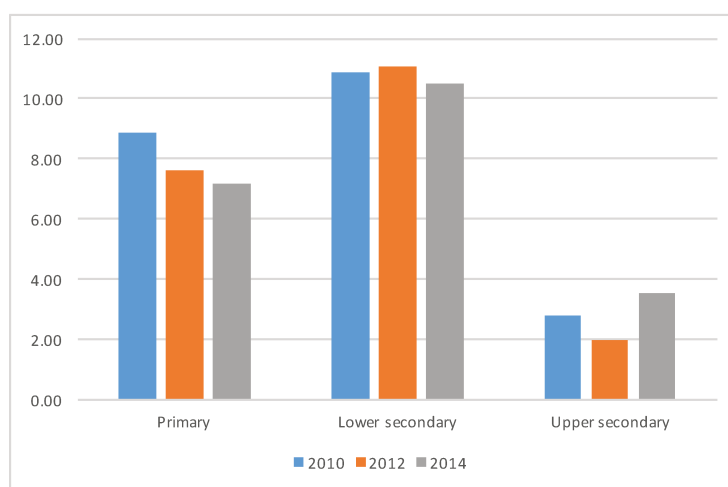
3.2. Performance of the education sector

Vietnam's education sector was thought to be performing well, given the economic conditions at the time. In 2014, Vietnam was ranked by the World Bank as 161th among 213 countries by gross national income per capita but ranked 116th out of 188 countries by human development index (HDI) - the upper end of the medium group according to the Vietnam Human Development Report 2015 (UNDP & VASS, 2016). The contribution of the improvements in education to the HDI growth was about 43 percent, which is even higher than that of the top human development group.

Vietnam has achieved impressive outcomes on the acquisition of cognitive skills. It ranked 12th out of 76 participating countries in mathematics, reading and science test for the 15-year-old students in the Programme for International Student Assessment (PISA) in 2012, which is above the OECD average as well as above the United States, Australia, France and the United Kingdom.

The Government declared that in 2008 practically all children in Vietnam had access to primary education. The net enrolment and completion rates for the primary level were over 98 percent and 92 percent in the school year 2012-2013 (MOET, 2014). However, it is not clear whether the universal lower secondary, another target of the Education Development Strategy for the period 2001–2010 and also of the National Education for All Action Plan 2003-2015, had been achieved or not by 2015. The net enrolment rate for lower secondary education has also increased significantly to 88 percent in 2012-2013 (from 70 percent in 2000), while the completion rate in this year was 83 percent. Own calculations using data from the Vietnam Household Living Standard Surveys in 2010, 2012 and 2014 show that the dropout rates of lower secondary schools for the 10 to 14-year-old adolescents nearly kept at 11 percent in these three years (Figure 1- 2). This figure was similar to MOET's dropout rate of 11.8 percent in 2009 (MOET, 2013).

Figure 1- 2: Dropout rates, by highest educational level completed (%)



Source: Author's calculation based on VHLSS 2010, 2012, 2014

4. Data and methodology

4.1. Data

The Young Lives Older Cohort completed primary schools mostly before Round 3 (97 percent) and the Younger Cohort mainly before Round 4 (89 percent). The figures mean that dropping out should occur starting from Round 3 – at the age of 15 for the Older Cohort and Round 4 – at the age of 12 for the Younger Cohort at the earliest. In fact, the adolescents from the Older Cohort not enrolling in any lower secondary grade account for 98 percent of the total not-enrolled in the level. As a result, this study relies on the panel data of the Older Cohort.

4.2. Dropouts

Dropping out can be specified based on the section on schooling in each round of the survey. In this study, the section on child's educational history in the child questionnaire in 2014 (Round 4) is of particular importance in determining when the children dropped out during the time span of the Young Lives project. The section records the adolescents' retrospective information on their enrolment for each school year since the first round of the survey till 2003, making a total of 13 intervals of time recorded (yearly) to take into account for each student. With the expectation of a more reliable definition of school dropout (due to much more intervals - 13 versus four intervals), the study thus relies on the retrospective data collected in the last round rather than the section on schooling for each round of survey. For each year, a dropout of lower secondary school is a child that reported that he or she (1) had completed any grade between Grade 6 and Grade 8; and (2) had enrolled in school in the previous year to the year of reference; and (3) did not enrolled in school in

the year of reference. The status of enrolment then is cross-checked using the sections on schooling in the same questionnaires across rounds. Children with missing information on whether they attended schools or not, or on highest grade/level completed (accounting for 10.7 percent of the sample), are excluded from the panel data. In total, about 8.5 percent of the sample left their lower secondary schools.

Table 1- 1: Number of dropouts by highest grade completed and academic year

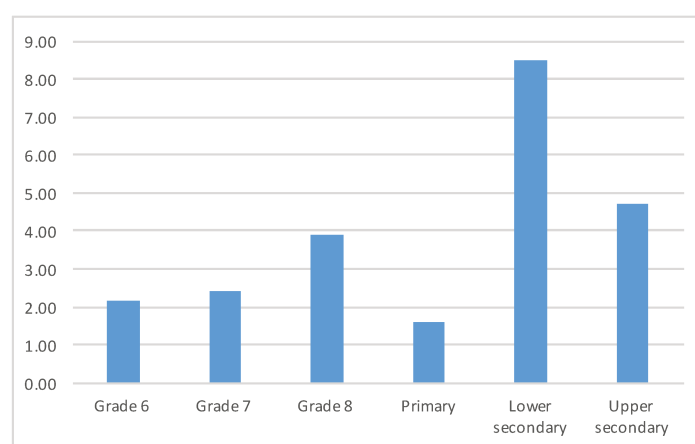
School year	Dropouts (No.)				Dropouts (%)			
	G6	G7	G8	Total	G6	G7	G8	Total
2000-2001	0	0	0	0	0.00	0.00	0.00	0.00
2001-2002	0	0	0	0	0.00	0.00	0.00	0.00
2002-2003	0	0	0	0	0.00	0.00	0.00	0.00
2003-2004	0	0	0	0	0.00	0.00	0.00	0.00
2004-2005	0	0	0	0	0.00	0.00	0.00	0.00
2005-2006	3	0	0	3	0.34	0.00	0.00	0.34
2006-2007	6	0	0	6	0.69	0.00	0.00	0.69
2007-2008	7	10	1	18	0.80	1.15	0.11	2.06
2008-2009	2	7	13	22	0.23	0.80	1.49	2.52
2009-2010	1	2	16	19	0.11	0.23	1.83	2.18
2010-2011	0	1	3	4	0.00	0.11	0.34	0.46
2011-2012	0	1	1	2	0.00	0.11	0.11	0.23
2012-2013	0	0	0	0	0.00	0.00	0.00	0.00
2013-2014	0	0	0	0	0.00	0.00	0.00	0.00

Source: Author's calculation from the Young Lives in Vietnam

Table 1- 1 provides the number of dropouts for lower secondary education across academic year and Figure 1- 3 and Figure 1- 4 depict the composition of dropouts by highest educational level completed and by academic year. Similar to the results from the representative data (Figure 1- 1), the share of Young Lives children who were dropouts at the lower secondary level here predominates those at both primary and upper secondary level. The proportion of dropouts of lower secondary schools, which was 8.5 percent, were five times and doubled the ones of the dropouts of primary and upper secondary schools. Dropping out was being observed from one year before Round 2 to two years before Round 4), and the phenomenon mostly occurred during 2007-2008 to 2009-2010. The dropout rates at lower secondary education increased over the years to reach its peak in 2008-2009 (one year before the year of Round 3, when the YL children aged 13-14), then decreased. Within the lower secondary level, dropout rates increased across grades. The proportions of children left school after completing Grade 6, and Grade 7 were highest in the school year 2007-2008, while the dropouts that completed Grade 8 gathered in 2009-2010, the time of Round 3 (Figure 1- 3).

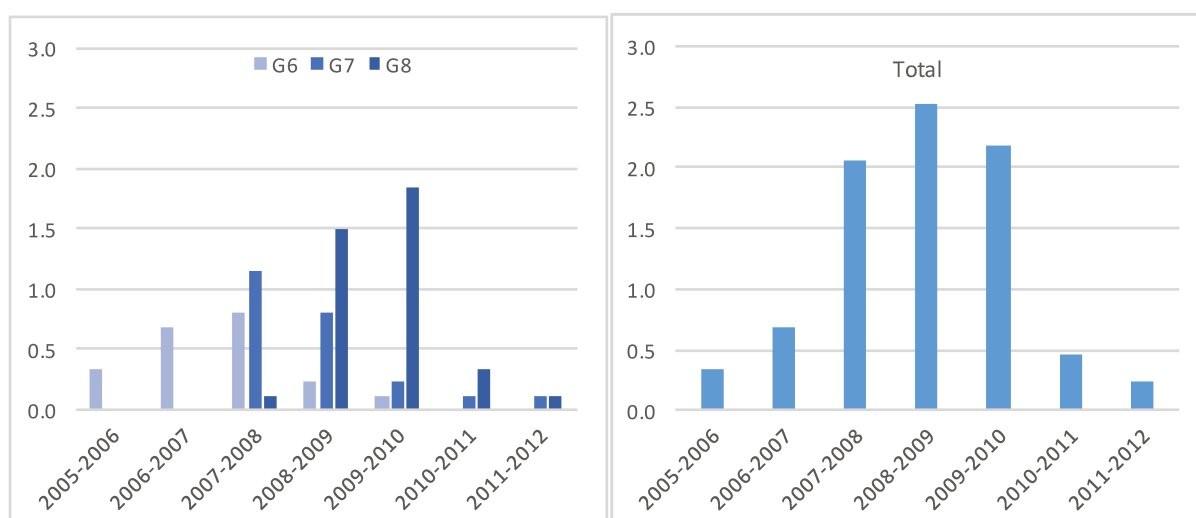
For the econometric analysis of the determinant of school dropouts, in addition to the schooling section, the paper mainly uses data in Round 2 and Round 3, the time when the event of dropping out occurred. The data drawn from the results of the tests (for Mathematics and Language/Vietnamese) administered to children, the responses to the questionnaire in the section on child's feeling and attitude and selected family characteristics. Because the data is observed four times (while the enrolment data is available for 13 years), the missing data for unmatched years before the year of each round of survey are filled with the data recorded for the previous round. It is a strong assumption because it supposes that skills evolve during the period between different rounds of the survey. However, the survival analysis of dropouts based on years involves more points of time (7 years) than on only two rounds, thus provides more reliable results. Because we look at the dropouts of lower secondary school, data for estimation excludes the children who had dropped out before entering a lower secondary level.

Figure 1- 3: Dropout rates, by highest lower secondary grade/level completed (%)



Source: Author's calculation from the Young Lives in Vietnam

Figure 1- 4: Dropout rates, by highest lower secondary grade completed and academic year (%)



Source: Author's calculation from the Young Lives in Vietnam

4.3. Measures of cognitive and non-cognitive skills

Because dropping out of lower secondary school was observed before the Young Lives Round 4, and scales of skills were administered starting from Round 2, it is possible to take into consideration the measures skills that are available in the second and third rounds of the survey.

In this chapter, measures of cognitive and non-cognitive skills are both single and composite. The single measures of skills include Rasch scores in Math, Peabody Picture Vocabulary Test (PPVT) for cognitive skill, and Self-esteem, Self-efficacy and Self-respect and Inclusion for non-cognitive skill. Definitions, computation and reliability of these measures are provided in Section 6.2 of the General Introduction. Briefly recalling, Self-efficacy is defined as an individual's belief in his/her capabilities to produce given attainments and to cope with adversity. An example of the statements related to self-efficacy is 'If I study hard at school I will be rewarded by a better job in future'. Self-esteem is how a person values or likes his self. One example of the statements related to pride is 'I am ashamed/proud of my shoes'. Self-respect and inclusion emphasise on the social component of self-esteem. The statements used to construct self-respect revolve around the concepts of pride and sense of inclusion, for instance, 'The other children in my class treat me with respect'. These non-cognitive skills are all documented as factors of dropping out of high school (see literature review on determinants of school leaving above).

The composite measures of cognitive skill are estimated using factor analysis of the single measures of skills, i.e. Math and Language Rasch scores for the composite cognitive skill in each round. Factor analyses show that the first components explain 77 and 78 percent of the variation in the composite cognitive skills in Round 2 and Round 3 respectively (see Appendix A.2). The high proportions justify the use of the first components as composite measures of cognitive skills in both rounds of the survey. The common factors also explain 77 to 80 percent of the variance in each single measures of cognitive skills.

Each composite measure of non-cognitive skill for each round is the average of normalised responses for all the statements used to construct all single measures of non-cognitive skills (Self-esteem, Self-efficacy and Inclusion) in the round. The composite measures are reliable given their values of Cronbach alpha of 0.74 and 0.72 for Round 2 and 3 respectively.

Variation in adolescent's endowment, inclusive of their cognitive and non-cognitive skills, across rounds of the survey (see Table 1- 4 below) justifies the use of Cox models in examining the determinants of dropping out in this chapter.

4.4. Estimation strategy

This chapter applies the Cox proportional hazard model to investigate the determinants of school leaving for lower secondary education. Cox model is the most popular mathematic modelling approach of survival analysis that takes into consideration several explanatory variables simultaneously. The methodology, as previously discussed in the literature review on survival analysis, can overcome the limitations of cross-section designs in dropout studies. By investigating the transition from being enrolled to not enrolled, the Cox model helps to consider not only whether or

not, but also when and why students leave school. The model allows dynamic modelling of dropping out, taking into account the problem of censoring. The semiparametric specification or unspecified function of baseline hazard also makes the Cox model mathematically attractive to researchers (see more details below).

Consider a student that is characterised by a set of variables x . The student will either drop out of school or not during their lower secondary level. It means that we can observe either the time to “failure”, which is dropping out, denoted by T , or not dropping out till the end of the data (censored). T is discrete because whether students drop out or not is observed annually (each school year). The standard Cox model following (Cox, 1972) is specified as follows:

$$\lambda(t|x) = \exp(x\beta) \cdot \lambda_0(t) \quad (1)$$

where $\lambda(t)$ is the hazard rate - the rate of failure at time t , i.e. the proportion of the students in year t that drop out between year t and $t+1$ or $\lambda(t) = P[t \leq T \leq t+1 | T \geq t]$; $x(t)$ denotes the time-varying explanatory variables, β is a vector of unknown parameters associated with x and $\lambda_0(t)$ is the baseline hazard – an unspecified function giving the hazard function of the standard set of condition $x(t) = 0$. The coefficients of the model are obtained using maximum likelihood estimation. For a survival analysis like the Cox model, the measures of effects are hazard ratios, which are expressed by the exponential of the estimated coefficients (which is different from the linear or logistic model where measures of effects are estimated coefficients or odds ratios respectively). Specifically, exponentiated individual coefficient e^{β_i} is the ratio of the hazards for a one-unit change in the corresponding covariate x_i (x^i are the same as x except that x_i changes by one unit)

$$\hat{\lambda}_i = \frac{\hat{\lambda}(t|x^i)}{\hat{\lambda}(t|x)} = \exp(\hat{\beta}(x^i - x)) \quad (2)$$

Once the model is fitted, the estimated hazard ratio is a constant that allows interpreting the effect independent of time. This time independence of the Cox model is thus a diagnostic to verify that a specification is not violating the Cox proportional hazard assumption.

Assuming that x including one vector of time-constant variables x_c and one vector on time-varying variable x_t , (1) can be expressed as:

$$\lambda(t|x_c, x_t) = \exp(x_c\beta_c + x_t\beta_t) \lambda_0(t) \quad (3)$$

If x_t comprise one vector of time-varying variables x_{tv} and time-varying covariates $x_{tc}(t)$ (variables whose effects change over time), (3) can be extended to incorporate a vector of interactions between the time-varying covariates and time variable $x_{tv}(t) = x_{tv} \times t$ to examine the whether the effects of these variables change over time the as follows:

$$\lambda(t|x_c, x_{tv}, x_{tc}(t)) = \exp(x_c\alpha_c + x_t\alpha_t + x_v(t)\alpha_v) \lambda_0(t) \quad (4)$$

The estimated hazard ratio obtained from the extended Cox model (4) is not a constant but is dependent on time.

The Cox proportional hazard model is widely used in empirical studies for its semiparametric specification. Although its functional form is not fully specified, the model is a robust one that provides reasonably good estimates of the coefficients and hazard ratios of interest. Mathematically, the Cox

model is also appealing because (i) the exponential form of the function ensures that the fitted hazard is non-negative; (ii) estimation of coefficients is possible without specifying hazard function; (iii) computation of hazard ratio is possible without estimating the baseline hazard function; (iv) estimation of the hazard and survival function is possible using minimum assumptions.

The explanatory variables x_c and x_t incorporated in the Cox models in this chapter are defined in Table 1- 2. Time constant variables x_c include adolescent's gender, mother's ethnicity, parental education and parental aspiration for their child's education (observed in Round 2). Time-varying variables x_t adolescent's educational aspiration, cognitive and non-cognitive skills expressed either in single or composite measures, time allocated for supporting the family such as doing household chores, the domestic task on household's farming/business, and paid activities, their household's wealth index, household size and location in residence. The wealth index and household size observed in the previous rounds are used instead of the variables observed in the same round to avoid the multicollinearity due to high correlation with between contemporaneous variables on family background. As regards residential location, the Young Lives Survey was conducted in five provinces (Lao Cai, Hung Yen, Da Nang, Phu Yen, Ben Tre) and among which Da Nang is the only urban region. Counting the panel data, Da Nang accounted for 98 percent of the adolescents living in urban areas, and about 1 percent of the sample has moved to other provinces (non-Young Lives provinces) in Round 3. Therefore the variables on residential location included in the model can be either the dummy (on the urban area or not, which is Da Nang) or four dummies on provinces (Hung Yen, Da Nang, Phu Yen, Ben Tre, excluding Da Nang as the base category). Given the negligible change in the residential location, these variables are the only group of time-varying variables being excluded from the time-varying covariates (i.e. not taken into account as the factors with varying effects).

Table 1- 2: Variables included in the models

Variable	Scale
<i>Time-constant variables</i>	
Male	0 = No; 1 = Yes
Mother's minority	0 = No; 1 = Yes
Caregiver's aspiration for child's education (Round 2)	1 = Less than university ³⁴ ; 2 = University
Father's education	1 = Below primary; 2 = Primary; 3 = Lower secondary; 4 = Upper secondary and above
Mother's education	1 = Below primary; 2 = Primary; 3 = Lower secondary; 4 = Upper secondary and above
<i>Time-varying variables</i>	
Urban (Da Nang)	Observed in Round 2 and Round 3 0 = No; 1= Yes

³⁴ "Less than university" includes such categories as "None", "Grade 1" to "Grade 12", "Adult literacy" and "Religious education"

Da Nang (Base category)	0 = No; 1= Yes
Lao Cai	0 = No; 1= Yes
Hung Yen	0 = No; 1= Yes
Phu Yen	0 = No; 1= Yes
Ben Tre	0 = No; 1= Yes
<i>Time-varying covariates</i>	<i>Observed in Round 2 and Round 3, if not otherwise stated</i>
Math	Normalized Rasch score
PPVT	Normalized Rasch score
Self-esteem	Average of normalized responses for all the statements used to construct the index
Self-efficacy	Average of normalized responses for all the statements used to construct the index
Inclusion	Average of normalized responses for all the statements used to construct the index
Cognitive skill	Predicted score from the factor analysis of the Rasch scores in Math and PPVT
Non-cognitive skill	Average of normalized responses for all the statements used to construct Self-esteem, Self-efficacy and Inclusion
Educational aspiration	1 = Less than university ³⁵ ; 2 = University; 3 =Master/Doctoral at university
Time spent on chores	Number of hours spent on household's chores per day
Time spent on domestic tasks	Number of hours spent on household's domestic task (farming, business) per day
Time spent on paid activity	Number of hours spent on paid activity per day
Prior wealth index	Average of the three indexes, including consumer durable, housing quality, access to services (Round 1 and Round 2)
Prior household size	Number of members in the household in Round 2 (Round 1 and Round 2)

Source: Young Lives Survey in Vietnam

Nearly all the variables included in the model (exclusive of adolescent's gender, Inclusion and urban areas) are significant to the hazard of dropping out (see results of the tests for the significance of variables in A.3). Variables on family's socioeconomic status (minority, wealth index and parental education) are highly correlated with each other and significantly associated with the variables on adolescent's individual and composite skills (see Appendix A.1). The correlations are as expected given the evidence on factors of dropping out (provided in the literature review above) as well as the relationship between family background and the extent to which parent devote time and care to their children during the skill development. Even so, the multidimensional approach in analysing the determinants of dropping out entails the consideration for as many factors as possible. For robustness

³⁵ "Less than university" includes "None", "Grade 1" to "Grade 12", "Incomplete technical/pedagogical institute", "Complete technical/pedagogical institute", "Incomplete University", "Others"

check, different Cox models are estimated, beginning with the inclusion of skills only, then adding in turn individual, family and community characteristics.

Testing for the proportional hazards assumption of a Cox model is also the test for a proper model specification. Standard methods to verify the Cox proportional hazard assumption of a model specification involve the test based on the re-estimation of the model and the test based on Schoenfeld residuals (Cleves et al., 2004). For the former, the most straightforward way is the link test as follows. Given the standard Cox model fit by (1) corresponds to the log relative hazards:

$$H = x\beta \quad (2)$$

The basis for the link test is estimating β from (2) and re-estimating β from $H = x\hat{\beta} + (x\hat{\beta})^2$. If $x\beta$ is the correct specification, then the coefficient on the squared linear predictor is insignificant. The other way is to estimate (2) adding the interactions between individual predictors and time (in other words, the time-varying covariates): $H = x\beta + (x\beta)t$. The insignificant of the coefficients on the interactions indicate that the model does not violate the assumption of proportionality.

The test based on Schoenfeld residuals entails the test of the nonzero slope in a linear regression of the scaled Schoenfeld residuals on some function of time. Given the Schoenfeld residual r_u for the failed observation j is the difference between the covariate value x_u for j and the weighted average of the covariate values of all observations at risk of failure when j failed:

$$r_u = x_u - \frac{\sum_{R_j} x_u \exp(x\hat{\beta})}{\exp(x\hat{\beta})}$$

the basis of the test is allowing x_u varies with time, or $\beta(t) = \beta + p_j q(t)$, where $q(t)$ is a function of time, then prove that the coefficient of the time function is insignificant ($p_j = 0$).

The goodness of fit test for the Cox model relies on the use of Cox-Snell residuals, which are specified as (\hat{H}_0 and $\hat{\beta}$ obtained from the Cox model):

$$CSr_j = \hat{H}_j(t_j) \exp(x_j \hat{\beta})$$

If the model is correctly defined, then the Cox-Snell residuals follow a standard exponential distribution with value 1 of the hazard function at all points in time, meaning the cumulative hazard of the Cox-Snell residuals is a straight 45° line. The fit of the model can be proved by the estimation of Nelson-Aalen cumulative hazard function, where time variable is the Cox-Snell residuals and variable on initial failure is the failure indicator. The more the plot of function against Cox-Snell residuals fit the 45° line, the more the Cox model fit the data.

5. Determinants of dropping out in Vietnam: empirical results

In total, 780 adolescents (of which dropouts account for nearly 7 percent) with non-missing values for all variables, make up the sample available for model estimation. Of the sample, female adolescents account for a higher share. The group from the ethnic majority and leaving in rural areas predominate at the proportion of 73 and 81 percent. Most of their parents – 77 percent of the fathers and 84 percent of the mothers obtained lower secondary level and less. The non-dropouts achieved

higher academic performance than their counterpart, signalling a lower level of skills of the dropouts compared with the non-dropouts. The differences in the skill measures are observed for all skills, except for Inclusion. It is notable that the dropouts devoted less time on the family's chores than the non-dropouts.

Table 1- 3: Description of time-constant variables

	All	Non-dropouts	Dropouts	
Male	0.46	0.46	0.55	**
Minority	0.13	0.13	0.22	***
Caregiver's aspiration for child's education (R2)	1.75	1.78	1.42	***
Father completed less than primary level	0.15	0.13	0.36	***
Father completed primary level	0.26	0.25	0.40	***
Father completed lower secondary level	0.36	0.37	0.22	***
Father completed at least upper secondary level	0.23	0.25	0.02	***
Mother completed less than primary level	0.21	0.18	0.58	***
Mother completed primary level	0.30	0.30	0.27	
Mother completed lower secondary level	0.34	0.35	0.11	***
Mother completed at least upper secondary level	0.16	0.17	0.04	***

Note: difference significant at 1% (***), 5% (**), 10% (*)

Source: Author's calculation from the Young Lives in Vietnam

Table 1- 3 and Table 1- 4 imply significant gaps regarding family background, skills and educational aspiration between the dropouts and non-dropouts. The dropouts were more often male, from the minorities, and had lower socioeconomic status than the average. While the aggregate shares of fathers and mothers with lower secondary level were 37 and 35 percent respectively, the figures for the dropouts were only 22 and 11 percent for the dropouts correspondingly. In a worse situation, up to 76 percent of the dropouts' fathers completed primary level at the highest and 58 percent of their mothers finished less than primary level.

The dropouts also came from the families with lower wealth index and higher household size than the adolescents remaining in schools. The negligible difference is observed for the location of the dropouts and non-dropouts. The non-dropouts had higher educational aspiration and achieved higher skills than their counterparts. The differences in the skill measures are observed for all cognitive and non-cognitive skills, except for Inclusion. The differences are more significant for cognitive than non-cognitive skills. The dropouts spent less time on the family's chores than the non-dropouts. Not counting the distribution of residential location being nearly unchanged, the remaining time-varying

factors (skills, educational aspiration, time allocation, wealth index and household size) considerably varied between Round 2 and 3 (Table 1- 4).

Table 1- 4: Description of time varying variables

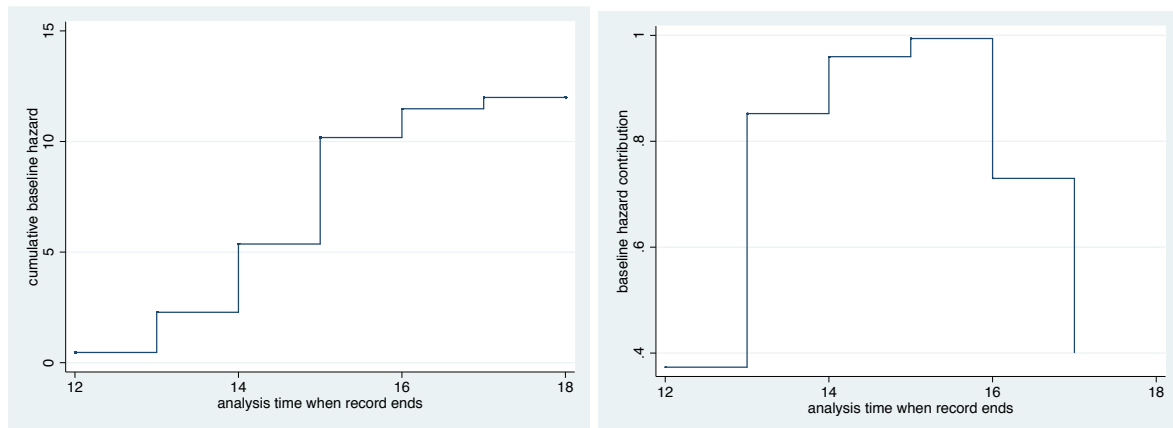
	Round 2				Round 3			
	All	Non-dropouts	Dropouts		All	Non-dropouts	Dropouts	
Math	0.03	0.07	-0.83	***	0.06	0.09	-0.89	***
PPVT	0.08	0.10	-0.38	***	-0.13	-0.10	-1.39	***
Self-efficacy	0.01	0.02	-0.09	*	0.03	0.03	-0.16	*
Self-esteem	0.01	0.02	-0.16	*	0.04	0.05	-0.10	*
Inclusion	0.01	0.00	0.17		0.04	0.05	-0.20	
Cognitive skill	0.03	0.07	-0.86	***	0.04	0.06	-0.84	***
Non-cognitive skill	0.01	0.01	-0.04	*	0.04	0.04	-0.14	*
Educational aspiration	1.77	1.78	1.44	***	1.73	1.74	1.24	***
Hours spent on household's chores	1.10	1.11	0.96	***	1.44	1.46	0.81	***
Hours spent on domestic tasks	0.59	0.59	0.56		0.94	0.89	2.62	
Hours spent on paid activities	0.03	0.02	0.09		0.35	0.32	1.67	
Lagged wealth index	0.52	0.52	0.44	***	0.64	0.64	0.47	***
Lagged household size	4.91	4.89	5.35	**	4.56	4.55	4.76	**
Urban	0.19	0.19	0.24		0.19	0.19	0.10	
Lao Cai	0.19	0.20	0.06		0.19	0.19	0.24	
Hung Yen	0.21	0.21	0.06	***	0.21	0.21	0.00	***
Da Nang	0.19	0.19	0.21		0.19	0.19	0.10	
Phu Yen	0.20	0.19	0.47	**	0.20	0.20	0.14	**
Ben Tre	0.21	0.21	0.18		0.21	0.20	0.52	

Note: difference significant at 1% (***), 5% (**), 10%(*)

Source: Author's calculation from the Young Lives in Vietnam

Figure 1- 5 provides the total number of dropouts that would occur up until each age, or the cumulative hazards in the left plot, and the rate of changes in the magnitude of the hazards – the hazard contribution in the right plot at which all covariates are zero (all explanatory variables are set to be zero). There were increases in hazard rate before age 16, where the hazards started to decline. The age of 13 and 17 are observed with highest increasing rate and decreasing rate of hazards.

Figure 1- 5: Baseline cumulative hazard and hazard contribution



Source: author's calculation based on the Young Lives in Vietnam

Table 1- 5 presents the estimation results of the Cox models in which each of the single skills (Model (1) to (5)), set of single cognitive and non-cognitive skills (Model (6) and (7)) and composite skills (Model (8) and (9)) are the only predictors of dropping out. The tables report the ratios of the hazards for one unit change in the corresponding skills. All single measures of skills except Inclusion are significant to the hazard. One standard deviation (SD) increase in Math would, *ceteris paribus*, decrease the hazard by 60 percent. For PPVT, self-esteem and self-efficacy, the decreases in hazards for one additional SD were smaller: the hazards reduced by 48, 40 and 36 percent respectively. Of the single skills that are examined separately, Math is the most significant and followed by PPVT, self-esteem and self-efficacy. When being considered together with other single non-cognitive skills in the model, the magnitudes of the predictors decline, and self-efficacy becomes insignificant. So Math is the most robust and significant predictor, then come PPVT and self-esteem. These predictors confirm the more significant effect of the composite cognitive skill compared with non-cognitive skill on the hazard of dropping out.

Table 1- 6 and Table 1- 7 show hazards ratios of predictors in the Cox models including single and composite measures of skills respectively (for robustness check of the model specifications, see Appendix A.6 and A.7 on the estimation results of the models incorporating the explanatory variables one by one). In both tables, the Cox models from left to right include factors of dropping out in turn, initially skills alone, then adolescent's characteristics, proxies for family and community's investment in adolescents. Column (13) and (14) in provide the results for full models that take into account single measures of skills, and column (20) and (21) in Table 1- 7 for the full models with composite skills. The correspondingly extended models with time-varying covariates are shown in column (15) and (16) in and column (22) and (23) Table 1- 7. In these columns for extended models, the left panel provides the hazard ratios for time-constant and time-varying variables, and the right panel for time-varying covariates.

Table 1- 5: Cox models with single and composite measures of cognitive and non-cognitive skills

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Math	0.3986*** (0.0508)					0.4918*** (0.0769)			
PPVT		0.5178*** (0.0530)				0.7347** (0.1019)			
Self-efficacy			0.6359** (0.1425)				0.6893 (0.1639)		
Self-esteem				0.6078** (0.1305)			0.5985** (0.1400)		
Inclusion					1.0046 (0.2209)		1.3537 (0.3365)		
Cognitive skill								0.4529*** (0.0489)	
Non-cognitive									0.5566* (0.1682)
No. of obs.	4,854	4,854	4,854	4,854	4,854	4,854	4,854	4,854	4,854
No. of failures	53	53	53	53	53	53	53	53	53
No. of subjects	780	780	780	780	780	780	780	780	780
Log-likelihood	-322.70	-330.73	-344.72	-344.06	-346.60	-320.38	-342.37	-325.02	-344.77
Chi2 p-value	0.0000	0.0000	0.0526	0.0243	0.9834	0.0000	0.0375	0.0000	0.0554
Adjusted R2	0.069	0.046	0.005	0.007	0.000	0.076	0.012	0.062	0.005

Note: *** p<0.01, ** p<0.05, * p<0.1

Source: author's calculation based on the Young Lives in Vietnam

The proportional hazard assumption of the models is verified using the tests based on Schoenfeld residuals and the tests based on the estimation of (14) and (20) in which same factors are included in both time-varying variables and time-varying covariates. For the former, the global test statistics of all models, and the test statistics of all predictors are insignificant, exclusive of Inclusion in (12) and (14). The test statistics of the latter show insignificant time-varying covariates. The test results together imply that the models do not violate for the proportional hazard assumption (see Appendix A.4 and A.5 respectively). The model specification for (14) and (21), the Cox models with time-varying covariates, is relatively good based on the goodness for fit test result provided in Appendix B.1 and B.2 correspondingly.

Table 1- 6: Cox models with single cognitive and non-cognitive skills

	(10)	(11)	(12)	(13)	(14)	(15)	(16)
	Without time-varying covariates					With time-varying covariates	
Math	0.5058*** (0.0792)	0.5756*** (0.0942)	0.5625*** (0.0977)	0.5613*** (0.0974)	0.5817*** (0.1011)	0.9316*** (0.0202)	0.9353*** (0.0203)
PPVT	0.7445** (0.1037)	0.7995 (0.1207)	1.0356 (0.1561)	1.0142 (0.1551)	1.0077 (0.1572)	1.0026 (0.0185)	1.0011 (0.0187)
Self-efficacy	0.7682 (0.1848)	0.8820 (0.2089)	0.9689 (0.2399)	0.9437 (0.2375)	0.9468 (0.2395)	0.9913 (0.0315)	0.9931 (0.0315)
Self-esteem	0.7157 (0.1739)	0.7156 (0.1771)	0.7272 (0.1772)	0.7358 (0.1802)	0.7469 (0.1804)	0.9652 (0.0290)	0.9673 (0.0286)
Inclusion	1.2835 (0.3179)	1.3710 (0.3314)	1.4393 (0.3499)	1.5110* (0.3787)	1.3837 (0.3552)	1.0414 (0.0314)	1.0307 (0.0318)
Educational aspiration		0.3263*** (0.1056)	0.5133** (0.1745)	0.5210* (0.1780)	0.5361* (0.1859)	0.9255* (0.0395)	0.9294* (0.0404)
Male		1.3322 (0.3762)	1.0703 (0.3110)	1.0248 (0.2986)	1.0058 (0.2947)	1.0447 (0.3045)	1.0165 (0.2981)
Minority		0.7600 (0.3056)	0.3207** (0.1537)	0.3085** (0.1522)	0.5005 (0.2773)	0.3167** (0.1585)	0.5163 (0.2866)
Father's education			0.6975* (0.1441)	0.6785* (0.1398)	0.6418** (0.1334)	0.6843* (0.1408)	0.6422** (0.1331)
Mother's education			0.6237** (0.1479)	0.6403* (0.1500)	0.7401 (0.1800)	0.6335* (0.1482)	0.7452 (0.1813)
Caregiver's aspiration			0.7219 (0.2506)	0.7077 (0.2484)	0.7145 (0.2517)	0.7196 (0.2553)	0.7290 (0.2602)
Prior wealth index			0.4499	0.1733*	0.1736	0.7944*	0.7890*

			(0.4249)	(0.1783)	(0.1965)	(0.1025)	(0.1120)
Prior household size			1.2102**	1.1930**	1.1653*	1.0206*	1.0189*
			(0.1034)	(0.1019)	(0.0993)	(0.0115)	(0.0114)
Household's chores			0.6921**	0.7096**	0.7283*	0.9617**	0.9637*
			(0.1154)	(0.1157)	(0.1210)	(0.0191)	(0.0195)
Domestic tasks			1.0903	1.1125	1.1722**	1.0096	1.0161*
			(0.0807)	(0.0834)	(0.0898)	(0.0087)	(0.0089)
Paid activities			1.0549	1.0618	1.0685	1.0053	1.0059
			(0.0766)	(0.0775)	(0.0796)	(0.0082)	(0.0084)
Urban areas				2.8131**	0.1237	2.6952**	
				(1.2297)	(0.0911)	(1.1593)	
Lao Cai					0.1812***		0.1163***
					(0.1456)		(0.0860)
Hung Yen					0.3850**		0.1830**
					(0.1853)		(0.1466)
Phu Yen					0.4046**		0.4116*
					(0.2237)		(0.1942)
Ben Tre					-0.9048		0.4156
					(0.5529)		(0.2271)
No. of obs.	4,854	4,854	4,854	4,854	4,854	4,854	4,854
No. of failures	53	53	53.0000	53	53	53.0000	53.0000
No. of subjects	780	780	780.0000	780	780	780.0000	780.0000
Log-likelihood	-318.67	-311.37	-291.77	-289.23	-286.25	-290.51	-287.00
Chi2 p-value	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Adjusted R2	0.081	0.102	0.158	0.166	0.162	0.162	0.172

Note: *** p<0.01, ** p<0.05, * p<0.1

Table 1- 7: Cox models with composite cognitive and non-cognitive skills

	(17)	(18)	(19)	(20)	(21)	(22)	(23)
	Without time-varying covariates					With time-varying covariates	
Cognitive skill	0.4594*** (0.0502)	0.5575*** (0.0756)	0.6961** (0.1071)	0.6671** (0.1054)	0.6937** (0.1108)	0.9473*** (0.0189)	0.9515** (0.0193)
Non-cognitive skill	0.7094 (0.2230)	0.8916 (0.2837)	1.0669 (0.3310)	1.1003 (0.3507)	1.0204 (0.3242)	1.0057 (0.0393)	0.9987 (0.0388)
Educational aspiration		0.2999*** (0.0970)	0.5298* (0.1857)	0.5480* (0.1936)	0.5567 (0.1987)	0.9312 (0.0407)	0.9342 (0.0414)
Male		1.3120 (0.3674)	1.1782 (0.3374)	1.1282 (0.3237)	1.0632 (0.3085)	1.1358 (0.3261)	1.0670 (0.3098)
Minority		0.8421 (0.3318)	0.2935** (0.1453)	0.2714** (0.1392)	0.4739 (0.2706)	0.2707** (0.1414)	0.4692 (0.2687)
Father's education			0.7010* (0.1441)	0.6844* (0.1401)	0.6370** (0.1317)	0.6900* (0.1412)	0.6407** (0.1321)
Mother's education			0.6121** (0.1423)	0.6351** (0.1468)	0.7362 (0.1757)	0.6326** (0.1461)	0.7404 (0.1767)
Caregiver's aspiration			0.6325 (0.2227)	0.6189 (0.2204)	0.6187 (0.2204)	0.6373 (0.2283)	0.6347 (0.2277)
Prior wealth index			0.5798 (0.5538)	0.2233 (0.2311)	0.1906 (0.2153)	0.8202 (0.1054)	0.7976 (0.1119)
Prior household size			1.2131** (0.1042)	1.1942** (0.1027)	1.1557* (0.0994)	1.0209* (0.0114)	1.0179 (0.0114)
Household's chores			0.7562* (0.1199)	0.7709* (0.1196)	0.7768 (0.1231)	0.9697 (0.0183)	0.9697 (0.0187)

Domestic tasks			1.1139 (0.0802)	1.1395* (0.0835)	1.2000** (0.0910)	1.0115 (0.0085)	1.0182** (0.0088)
Paid activities			1.0743 (0.0759)	1.0807 (0.0771)	1.0814 (0.0787)	1.0069 (0.0081)	1.0069 (0.0082)
Urban areas				2.8877** (1.2617)		2.8033** (1.2041)	
Lao Cai					0.1068*** (0.0795)		0.1005*** (0.0758)
Hung Yen					0.1683** (0.1345)		0.1713** (0.1366)
Phu Yen					0.3932** (0.1868)		0.4083* (0.1905)
Ben Tre					0.3660* (0.2019)		0.3684* (0.2001)
No. of obs.	4,854	4,854	4,854	4,854	4,854	4,854	4,854
No. of failures	53	53	53.0000	53	53	53	53
No. of subjects	780	780	780.0000	780	780	780	780
Log-likelihood	-324.43	-316.04	-296.47	-293.81	-290.25	-294.17	-290.27
Chi2 p-value	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Adjusted R2	0.064	0.088	0.145	0.152	0.163	0.151	0.163

Note: *** p<0.01, ** p<0.05, * p<0.1

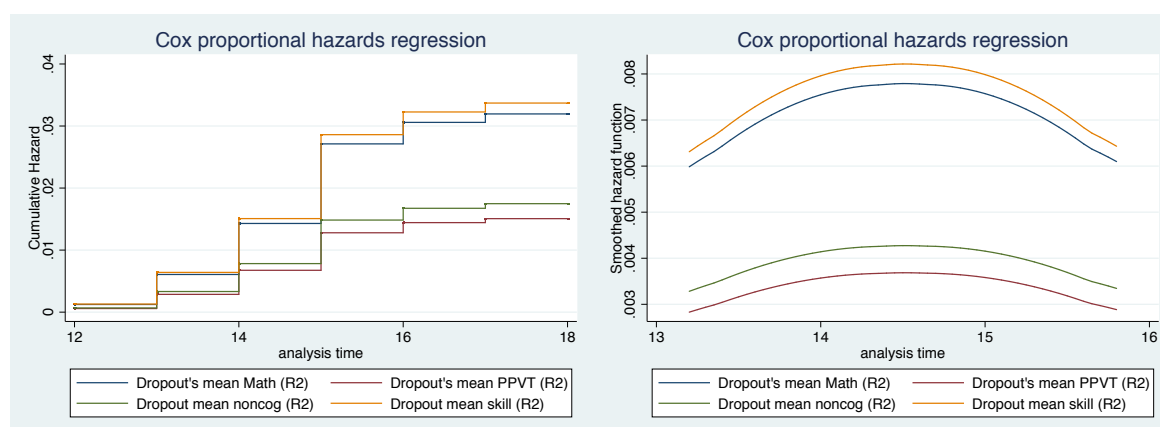
Source: Author's calculation from the Young Lives in Vietnam

Table 1- 6 reconfirms that Math is the most significant skill that predicts the hazard. One SD increase in Math, other things being equal, is associated with a reduction of 42 to 44 percent in the likelihood of dropping out. The remaining measures of skills are not robustly significant across models. PPVT is significant to a lesser extent only in Model (10) not controlling for other factors than skills. The ratio of hazards for Inclusion only in Model (13) implies that students with one more SD in Inclusion would, holding other things unchanged, face the one and a half more risk of school leaving. However, this counter-intuitive effect can be unreliable given the results of tests (see Appendix A.3, A.4 and A.5) that suggest Inclusion violates the proportional hazard assumption.

The effect of Math on the probability of dropping out at an earlier age magnify across later ages. For each passing year, one SD in Math averagely raises the contemporaneous hazard by 6 to 7 percent compared to the hazard in the previous year. Regarding composite skills, Table 1- 7 similarly shows the only cognitive skill emerges as the most significant determinant of dropping out. A student with one more SD in cognitive skill would, *ceteris paribus*, be 31 to 33 percent less likely to stop schooling. The protective role of cognitive skill is on average rising by 5 percent compared to one year earlier. Non-cognitive skill always has no predictive power whatever the measures of skills being used in the full models with all determinants of dropping out than only skills.

In addition to cognitive skill, especially Math, adolescent's educational aspiration has similar predictive power on the hazards. Other things being equal, each higher level of adolescent's aspiration is associated with 46 to 67 percent less in the probability of school leaving. For each passing year, the higher level of educational aspiration reduces the probability in the previous by 5 percent (model (22)-(23)) to 7 percent (model (15)-(16)). Caregiver's aspiration for adolescents' education is insignificant probably because it is absorbed in adolescent's educational aspiration through the care devoted as well as ambition for education transmitted to adolescents during their childhood and schooling.

Figure 1- 6: Cumulative hazards and hazard contribution, subgroups with skills equal to dropout's mean skills



Source: author's calculation based on the Young Lives in Vietnam

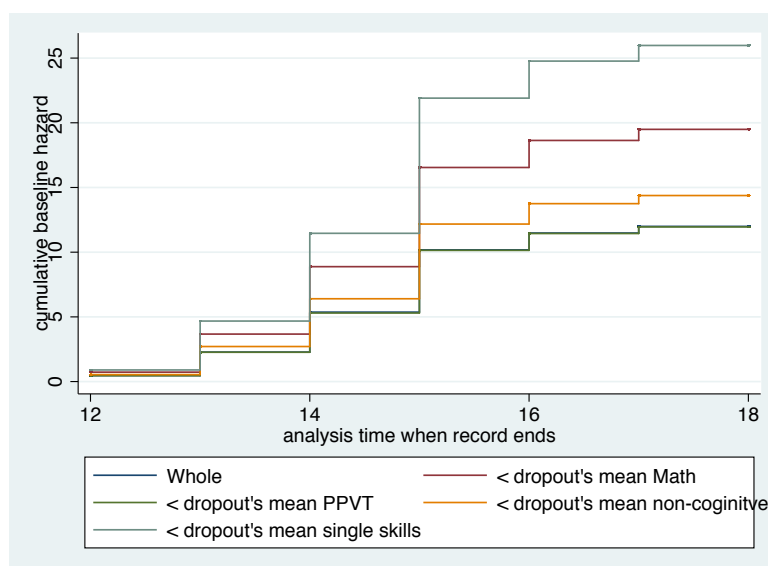
The cumulative hazards (left plot in Figure 1- 6) and increments of cumulative hazards (right plot in Figure 1- 6) at each event time of the non-dropouts obtaining same mean values of skill(s) as of the dropouts in Round 2 show that these non-dropouts faced higher hazards than the average of the whole sample. Of these non-dropouts, adolescent with all single skills (Math, PPVT, Self-efficacy and Self-esteem) equal to the dropouts' mean skills have the highest hazards, following by the hazards of those with same Math score, then of those with same self-efficacy and self-esteem (labelled as “noncog” in the figure) and those with same PPVT.

Table 1- 8: Baseline survival and hazard of selected groups

	Baseline survival	Baseline cumulative hazard	Baseline hazard contribution	Relative hazard
Single skills				
Whole	0.1102	7.4496	0.8753	0.0059
Less than dropout's mean Math R2	0.0767	12.1202	0.9358	0.0037
Less than dropout's mean PPVT R2	0.1110	7.4156	0.8736	0.0060
Less than dropout's mean self-efficacy & -esteem R2	0.0978	8.9151	0.9019	0.0050
Less than dropout's mean skills as above R2	0.0987	16.0234	0.9555	0.0028
Composite skills				
Whole	0.1058	7.6039	0.8785	0.0058
Less than dropout's mean cognitive skill R2	0.0643	14.6860	0.9514	0.0030
Less than dropout's mean non-cognitive skill R2	0.1006	8.1455	0.8895	0.0054
Less than cognitive & non-cognitive skill R2	0.0632	14.9775	0.9525	0.0030

Source: Author's calculation from the Young Lives in Vietnam

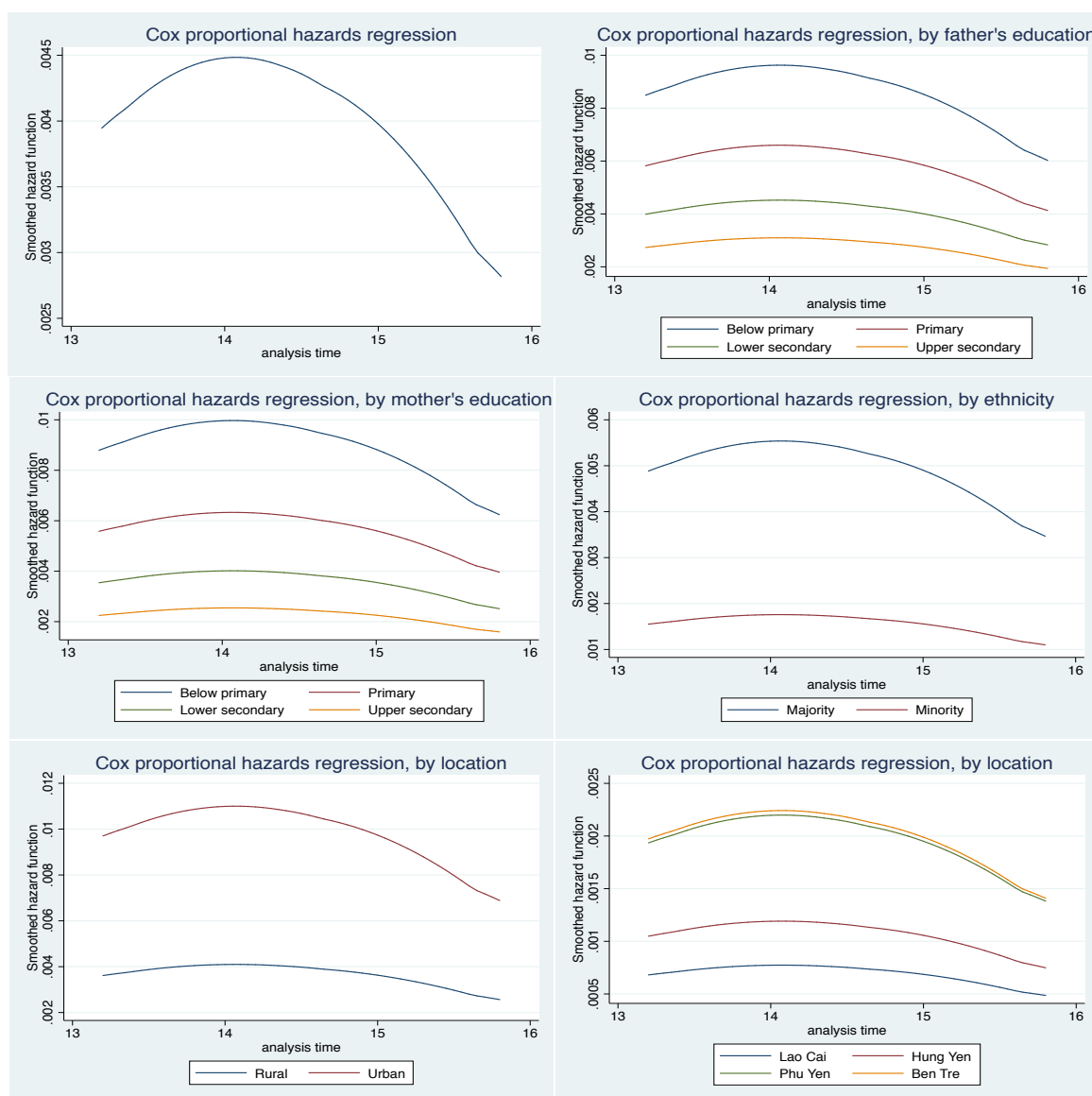
Figure 1- 7: Baseline cumulative hazards, subgroups with skills less than dropout's mean skills



Source: Author's calculation from the Young Lives in Vietnam

Similar trend is observed for the non-dropouts with lower skill(s) than the dropout's mean skill(s) (Table 1- 8). The non-dropouts with all single skills lower than the dropout's mean skill face highest risk of leaving school, then come the group with lower Math scores; lower self-efficacy and self-esteem and finally those with lower PPVT (Figure 1- 7).

Figure 1- 8: Estimated hazard functions, whole sample and sub-group



Source: Author's calculation based on the Young Lives in Vietnam

Ethnicity, parental education and residential location are the significant time-constant factors of dropping out (see Figure 1- 8 for the increments of the estimated cumulative hazards over time for the factors, Appendix B.4 for the cumulative hazards). Father's and mother's level of education have nearly equal predictive power. Other things remaining the same, adolescents with one additional level obtained by their father or mother would be 32 to 36 percent less likely to withdraw from school. Father's education seems to be more robustly significant predictor than mother's education. It is

noteworthy that ethnicity and location are significant at greater magnitudes than parental education, and minorities, as well as urban areas, are observed with lower hazards than their counterparts. Students of minorities have 68 to 73% less in the probability to drop out. Meanwhile, the likelihood of dropping out of urban students/in Da Nang is 2.7 to 2.9 times that of rural areas. The probabilities that rural students face the risk of dropping out in other provinces are only 11 to 41 percent of students in Da Nang.

6. Conclusions and policy discussion

Access to basic schooling, which includes primary and lower secondary level, is crucial to acquire foundation skills for further learning and livelihood in today's increasingly globalised world demand. Because the primary education is nearly universal worldwide while the global rate of out-of-school at lower secondary schooling age have risen since 2010, the dropping out at this educational level is a policy concern for many countries. Vietnam is not the exception with the net primary and lower secondary enrolment rate of 98 percent and 88 percent in 2012 and nearly constant dropout rate of 11 percent since 2010. The situation could lead the dropouts' disadvantages in employment and earnings in comparison to the graduates and the lack of skilled labour to serve the country's inclusive growth. Meanwhile, a few of existing evidence on why students leave school address either the nature of dropping out as a dynamic process or the significance of skills, particularly the non-cognitive type. Neither such study is available for Vietnam. This chapter, therefore, applies Cox proportional hazards models to examine the determinants of lower secondary school leaving as a dynamic process, taking into account both cognitive and non-cognitive skills. Though the Cox model does not identify causal causes of dropping out, given its strength in analysing dropping out as a dynamic and multidimensional process, the results can be reliable and provide useful hints for reducing the lower secondary dropout rate in the future.

The results of the goodness of fit tests prove the validity of the model specifications. Estimation results indicate that beginning around the age 13, the risk of dropping out was increasing to reach its peak at age 14, then decreasing. Among the factors, the most significant factors include the cognitive skill, parental education, adolescent's educational aspiration, and time allocated for household's chores. The effects of Math, adolescent's educational aspiration and socioeconomic status (household's wealth index, household size and time children spent on household's chores) on the probability of dropping out at an earlier age magnify across later ages.

The results imply the inequality of opportunity in remaining in lower secondary schools. Since dropping out is a multidimensional and dynamic process that starts early, the policy to restrain the phenomenon should cover many "factors". According to Rumberger (2011), the intervention could be designed either systematically to improve the environment of the students prone to drop out by providing supports on resources or other assistance. Or it could be developed as a program to improve the students' behaviours, feeling and perception of schooling. The tested policies also show the combination of both strategies.

Moreover, the analysis also emphasises that poor academic performance or low cognitive skill, particularly in Math, gives rise to adolescents' decision to drop out. A small lack of Math skill from the

beginning can be accumulated to push an adolescent in the risk of dropping out later on. The evidence suggests that observing and intervene to improve adolescent's skill since the early age is protective to school leaving. The measures to enhance academic achievement involves not only students but also other agents and efforts. As reviewed in De Witte et al. (2013), one of such measure could be introducing counselling programs that entail the participation of teachers, family members, peers and even community advisors in playing as role models, in supporting and motivating students in their academic performance. Because the involvement of parents in their children's performance at school is crucial, parent's supervision and care for their children could be supported through counselling program for parents or parents' discussion groups. Also, disadvantaged students can work during their high school but are not encouraged to get an adult job before they turn 16 years old (Entwisle et al., 2005). Finally, the intervention would be programs that assist disadvantaged parents (i.e. single parent, poor parents) so that they can ensure family's living conditions or environment that encourage children's studying.

Reference

- Al-Hroub, A. (2014). Perspectives of school dropouts' dilemma in Palestinian refugee camps in Lebanon: An ethnographic study. *International Journal of Educational Development*, 35, 53–66. <https://doi.org/10.1016/j.ijedudev.2013.04.004>
- Alivernini, F., & Lucidi, F. (2011). Relationship Between Social Context, Self-Efficacy, Motivation, Academic Achievement, and Intention to Drop Out of High School: A Longitudinal Study. *The Journal of Educational Research*, 104(4), 241–252. <https://doi.org/10.1080/00220671003728062>
- Allensworth, E. M. (2005). Dropout rates after high-stakes testing in elementary school: A study of the contradictory effects of Chicago's efforts to end social promotion. *Educational Evaluation and Policy Analysis*, 27(4), 341–364.
- Battin-Pearson, S., Newcomb, M. D., Abbott, R. D., Hill, K. G., Catalano, R. F., & Hawkins, J. D. (2000). Predictors of early high school dropout: A test of five theories. *Journal of Educational Psychology*, 92(3), 568–582. <https://doi.org/10.1037/0022-0663.92.3.568>
- Baulch, B., Vu, H. D., & Nguyen, T. (2012). Do Vietnamese Schools Provide the Right Education for an Industrialising Country? Young Lives Working Paper 81.
- Becker, G. S., & Tomes, N. (1979). An Equilibrium Theory of the Distribution of Income and Intergenerational Mobility. *Journal of Political Economy*, 87(6), 1153–1189.
- Becker, G. S., & Tomes, N. (1986). Human Capital and the Rise and Fall of Families. *Journal of Labor Economics*, 4(3, Part 2), S1–S39. <https://doi.org/10.1086/298118>
- Borgna, C., & Struffolino, E. (2017). Pushed or pulled? Girls and boys facing early school leaving risk in Italy. *Social Science Research*, 61, 298–313. <https://doi.org/10.1016/j.ssresearch.2016.06.021>
- Bradley, C. L., & Renzulli, L. A. (2011). The Complexity of Non-Completion: Being Pushed or Pulled to Drop Out of High School. *Social Forces*, 90(2), 521–545. <https://doi.org/10.1093/sf/sor003>
- Branson, N., Hofmeyr, C., & Lam, D. (2014). Progress through school and the determinants of school dropout in South Africa. *Development Southern Africa*, 31(1), 106–126. <https://doi.org/10.1080/0376835X.2013.853610>
- Bui, T. Q. (2011). School Dropout Trends in Vietnam from 1988 to 2006. In *Education in Vietnam* (J. London). Singapore: Institute of Southeast Asian Studies.
- Caprara, G. V., Fida, R., Vecchione, M., Del Bove, G., Vecchio, G. M., Barbaranelli, C., & Bandura, A. (2008). Longitudinal analysis of the role of perceived self-efficacy for self-regulated learning in academic continuance and achievement. *Journal of Educational Psychology*, 100(3), 525–534. <https://doi.org/10.1037/0022-0663.100.3.525>
- Cataldi, E. F., Laird, J., & KewalRamani, A. (2009). *High School Dropout and Completion Rates in the United States: 2007 (NCES 2009-064)*. National Center for Education Statistics, Institute of Education Sciences, US Department of Education. Retrieved from <http://files.eric.ed.gov/fulltext/ED506561.pdf>
- Cleves, M. A., Gould, W. W., & Gutierrez, R. G. (2004). *An introduction to survival analysis using stata* (Rev. ed). College Station, Tex: Stata Press.
- Coneus, K., Gernandt, J., & Saam, M. (2011). Noncognitive Skills, School Achievements and Educational Dropout. *Schmollers Jahrbuch*, 131(4), 547–568. <https://doi.org/10.3790/schm.131.4.547>

- Cooper, C. R., Chavira, G., & Mena, D. D. (2005). From Pipelines to Partnerships: A Synthesis of Research On How Diverse Families, Schools, and Communities Support Children's Pathways Through School. *Journal of Education for Students Placed at Risk (JESPAR)*, 10(4), 407–430. https://doi.org/10.1207/s15327671espr1004_4
- Cox, D. R. (1972). Regression Models and Life-Tables. *Journal of the Royal Statistical Society*, 34(2), 187–220.
- De Witte, K., Cabus, S., Thyssen, G., Groot, W., & van den Brink, H. M. (2013). A critical review of the literature on school dropout. *Educational Research Review*, 10, 13–28. <https://doi.org/10.1016/j.edurev.2013.05.002>
- DesJardins, S. L., Ahlburg, D. A., & McCall, B. P. (2006). The effects of interrupted enrollment on graduation from college: Racial, income, and ability differences. *Economics of Education Review*, 25(6), 575–590. <https://doi.org/10.1016/j.econedurev.2005.06.002>
- Entwisle, D. R., Alexander, K. L., & Steffel-Olson, L. (2005). Urban Teenagers: Work and Dropout. *Youth & Society*, 37(1), 3–32. <https://doi.org/10.1177/0044118X04268313>
- Finn, J. D. (1989). Withdrawing from school. *Review of Educational Research*, 59(2), 117–142.
- Fleisch, B., Shindler, J., & Perry, H. (2012). Who is out of school? Evidence from the Statistics South Africa Community Survey. *International Journal of Educational Development*, 32(4), 529–536. <https://doi.org/10.1016/j.ijedudev.2010.05.002>
- Glewwe, P. (2004). An Overview of Economic Growth and Household Welfare in Vietnam in the 1990s. In *Economic growth, poverty, and household welfare in Vietnam* (Paul Glewwe, Nisha Agrawal, and David Dollar). Washington, DC: World Bank.
- GSO, G. S. O., & UNICEF. (2014). Vietnam Multiple Indicator Cluster Survey 2014, Key Findings. Hanoi, Vietnam.
- Janosz, M., LeBlanc, M., Boulerice, B., & Tremblay, R. E. (1997). Disentangling the Weight of School Dropout Predictors: A Test on Two Longitudinal Samples. *Journal of Youth and Adolescence*, 26(6), 733–762.
- Kalmijn, M., & Kraaykamp, G. (2003). Dropout and Downward Mobility in the Educational Career: An Event-History Analysis of Ethnic Schooling Differences in the Netherlands. *Educational Research and Evaluation*, 9(3), 265–287. <https://doi.org/10.1076/edre.9.3.265.15572>
- Kaplan, D. S., Peck, B. M., & Kaplan, H. B. (1995). A structural model of dropout behavior: a longitudinal analysis. *Applied Behavioral Science Review*, 3(2), 177–193. [https://doi.org/10.1016/S1068-8595\(95\)80006-9](https://doi.org/10.1016/S1068-8595(95)80006-9)
- Kaufman, P., Alt, M. N., & Chapman, C. D. (2004). *Dropout rates in the United States: 2001 (NCES 2005-046)*. National Center for Education Statistics, US Department of Education. Retrieved from <http://files.eric.ed.gov/fulltext/ED483073.pdf>
- Korhonen, J., Linnanmäki, K., & Aunio, P. (2014). Learning difficulties, academic well-being and educational dropout: A person-centred approach. *Learning and Individual Differences*, 31, 1–10. <https://doi.org/10.1016/j.lindif.2013.12.011>
- Le, T. D., & Tran, N. M. T. (2013). Why Children in Vietnam Drop out of School and What They Do After That. Young Lives Working Paper 102.

- Liu, W. C., & Wang, C. K. J. (2008). Home Environment and Classroom Climate: An Investigation of their Relation to Students' Academic Self-Concept in a Streamed Setting. *Current Psychology*, 27(4), 242–256. <https://doi.org/10.1007/s12144-008-9037-7>
- Marks, G. N., Fleming, N., & Australian Council for Educational Research. (1999). *Early school leaving in Australia: findings from the 1995 year 9 LSAY cohort*. Camberwell, Vic.: Australian Council for Educational Research.
- Marsh, H. W., & Hau, K.-T. (2003). Big-Fish--Little-Pond effect on academic self-concept: A cross-cultural (26-country) test of the negative effects of academically selective schools. *American Psychologist*, 58(5), 364–376. <https://doi.org/10.1037/0003-066X.58.5.364>
- Ministry of Education and Training. (2013). All Children in School by 2015: Global Initiative on Out-of-School Children - Out-of-school Children in Vietnam: A country study. UNESCO Institute for Statistics and UNICEF.
- Ministry of Education and Training, (MOET). (2014). *Vietnam National Education For All 2015 Review*. Retrieved from <http://unesdoc.unesco.org/images/0023/002327/232770e.pdf>
- Nguyen, P. L. (2006). Effects of social class and school conditions on educational enrollment and achievement of boys and girls in rural Viet Nam. *International Journal of Educational Research*, 45(3), 153–175. <https://doi.org/10.1016/j.ijer.2006.11.008>
- No, F., Taniguchi, K., & Hirakawa, Y. (2016). School dropout at the basic education level in rural Cambodia: Identifying its causes through longitudinal survival analysis. *International Journal of Educational Development*, 49, 215–224. <https://doi.org/10.1016/j.ijedudev.2016.03.001>
- Paura, L., & Arhipova, I. (2014). Cause Analysis of Students' Dropout Rate in Higher Education Study Program. *Procedia - Social and Behavioral Sciences*, 109, 1282–1286. <https://doi.org/10.1016/j.sbspro.2013.12.625>
- Rumberger, R. W. (1983). Dropping out of High School: The Influence of Race, Sex, and Family Background. *American Educational Research Journal*, 20(2), 199–220.
- Rumberger, R. W. (1987). High School Dropouts: A Review of Issues and Evidence. *Review of Educational Research*, 57(2), 101–121.
- Rumberger, R. W. (2004). Why students drop out of school? In *Dropouts in America: confronting the graduation rate crisis* (G. Orfield, pp. 131–155). Cambridge, MA: Harvard Education Press.
- Rumberger, R. W. (2011). Dropping out: why students drop out of high school and what can be done about it. Cambridge, Mass: Harvard University Press.
- Sabates, R., Hossain, A., & Lewin, K. M. (2013). School drop out in Bangladesh: Insights using panel data. *International Journal of Educational Development*, 33(3), 225–232. <https://doi.org/10.1016/j.ijedudev.2012.09.007>
- Schunk, D. H., & Pajares, F. (2001). The development of academic self efficacy. In *Development of achievement motivation* (A. Wigfield & J. S. Eccles, pp. 16–29). San Diego, CA: Academic Press.
- Schunk, D. H., & Zimmerman, B. J. (2014). Competence and Control Beliefs: Distinguishing the Means and Ends. In *Handbook of Educational Psychology*. Routledge. <https://doi.org/10.4324/9780203874790.ch16>
- Sen, A. (2003). The importance of basic education: Amartya Sen's speech to the Commonwealth education conference, Edinburgh. Retrieved from <http://www.paddle.usp.ac.fj/collect/paddle/index/assoc/comsec3.dir/doc.pdf>

- Smeyers, P. (2007). The Relevance of Irrelevant Research; The Irrelevance of Relevant Research. In P. Smeyers & M. Depaepe (Eds.), *Educational Research: Why 'What Works' Doesn't Work* (pp. 95–108). Dordrecht: Springer Netherlands. https://doi.org/10.1007/978-1-4020-5308-5_6
- UIS & GEMR, (Institute for Statistics and Global Education Monitoring Report). (2016). Leaving no one behind: How far on the way to universal primary and secondary education? Policy Paper 37/ Fact sheet 27. Retrieved from <http://uis.unesco.org/sites/default/files/documents/fs37-leaving-no-one-behind-how-far-on-the-way-to-universal-primary-and-secondary-education-2016-en.pdf>
- UIS, (Institute for Statistics) (Ed.). (2011). *Comparing education statistics across the world*. Montreal, Quebec: UNESCO Institute for Statistics.
- UNESCO (Ed.). (2014). Teaching and learning: achieving quality for all. Paris: UNESCO.
- United Nations Development Programme (UNDP) and Viet Nam Academy of Social Sciences (VASS). (2016). *Vietnam Human Development Report 2015 on Inclusive Growth*. Hanoi, Vietnam: Social Sciences Publishing House.
- Vizcain, D. C. (2005). Investigating the hispanic/latino male dropout phenomenon: Using logistic regression and survival analysis. (Unpublished Doctoral Dissertation), Department of Educational Measurement and Research, College of Education, University of Florida. Retrieved from <http://scholarcommons.usf.edu/cgi/viewcontent.cgi?article=1895&context=etd>
- Vo, T. T., & Trinh, Q. L. (2005). Can Viet Nam Achieve One of Its MDGs? An Analysis of Schooling Dropout of Children. William Davidson Institute Working Paper 776, Ann Arbor: University of Michigan Business School.
- Willett, J. B., & Judith, S. D. (1991). From Whether to When: New Methods for Studying Student Dropout and Teacher Attrition. *Review of Educational Research*, 61(4), 407–450.
- Woldehanna, T., & Hagos, A. (2015). ECONOMIC SHOCKS AND CHILDREN'S DROPOUT FROM PRIMARY SCHOOL: IMPLICATIONS FOR EDUCATION POLICY IN ETHIOPIA. *Africa Education Review*, 12(1), 28–47. <https://doi.org/10.1080/18146627.2015.1036548>
- Yi, H., Zhang, L., Luo, R., Shi, Y., Mo, D., Chen, X., ... Rozelle, S. (2012). Dropping out: Why are students leaving junior high in China's poor rural areas? *International Journal of Educational Development*, 32(4), 555–563. <https://doi.org/10.1016/j.ijedudev.2011.09.002>

Appendix

A. Tables

A.1. Pairwise correlation between explanatory variables

	math	ppvt	self-esteem	self-efficacy	inclusion	cognitive skill	non-cognitive skill	child's aspiration	male	minority	dad's education	mom's education	caregiver's aspiration	lagged wealth index	lagged household size	chore	task	paid work	urban	lao cai	hung yen	phu yen	ben tre
math	1																						
ppvt	0.5264*	1																					
self-efficacy	0.1393*	0.1224*	1																				
self-esteem	0.1086*	0.1160*	0.3415*	1																			
inclusion	0.0630*	0.0485*	0.2713*	0.4314*	1																		
cognitive skill	0.8701*	0.8237*	0.1317*	0.1280*	0.0591*	1																	
non-cognitive skill	0.1228*	0.1301*	0.6604*	0.8473*	0.7081*	0.1304*	1																
child's aspiration	0.3514*	0.3271*	0.1781*	0.1192*	0.1424*	0.3802*	0.1722*	1															
male	-0.0634*	0.0128	-0.0265*	0.00320	-0.0219	-0.0348*	-0.0143	-0.1013*	1														
minority	-0.2334*	-0.3680*	-0.0785*	-0.0531*	-0.0582*	-0.2949*	-0.0743*	-0.2213*	0.00780	1													
dad's education	0.3250*	0.3530*	0.1487*	0.1154*	0.0590*	0.3626*	0.1355*	0.3224*	-0.0649*	-0.3957*	1												
mom's education	0.3653*	0.3823*	0.1263*	0.0983*	0.0495*	0.4009*	0.1163*	0.3218*	0.00320	-0.4177*	0.6581*	1											
caregiver's aspiration	0.3477*	0.3486*	0.1850*	0.1267*	0.0756*	0.3725*	0.1648*	0.4059*	-0.0615*	-0.2177*	0.3812*	0.3403*	1										
lagged wealth index	0.2854*	0.3884*	0.1148*	0.0840*	0.0201	0.3615*	0.0915*	0.2959*	0.0110	-0.4030*	0.5187*	0.5146*	0.3770*	1									
lagged household size	-0.1105*	-0.1777*	-0.00870	-0.0636*	-0.0322*	-0.1473*	-0.0549*	-0.1239*	-0.0643*	0.3170*	-0.1697*	-0.2479*	-0.1400*	-0.1105*	1								
chore	-0.0443*	-0.0680*	-0.00870	-0.0335*	0.00910	-0.0496*	-0.0214	-0.0221	-0.2259*	0.0384*	-0.0529*	-0.0743*	-0.0454*	-0.0467*	0.0185	1							
task	-0.2241*	-0.2952*	-0.0966*	-0.0858*	-0.0766*	-0.2433*	-0.1096*	-0.2429*	0.0518*	0.3710*	-0.2425*	-0.2542*	-0.1653*	-0.1953*	0.1873*	0.00260	1						
paid work	-0.1075*	-0.1518*	-0.0627*	-0.0604*	-0.0184	-0.1169*	-0.0702*	-0.0922*	0.0170	-0.0272*	-0.1077*	-0.0874*	-0.1330*	0.0363*	-0.0539*	-0.1045*	-0.0811*	1					
urban	0.1403*	0.2835*	0.0275	0.00180	-0.0613*	0.2003*	-0.0145	0.0726*	0.0399*	-0.1607*	0.2201*	0.2076*	0.1621*	0.4695*	0.0353*	-0.1189*	-0.1236*	0.000100	1				
lao cai	-0.1513*	-0.1789*	-0.0200	0.00950	-0.0150	-0.1444*	0.00510	-0.0903*	-0.0280*	0.5127*	-0.2328*	-0.1836*	-0.0942*	-0.3451*	0.1268*	0.0906*	0.2839*	-0.0510*	-0.2323*	1			
hung yen	0.0984*	0.0961*	0.0835*	-0.0110	-0.0161	0.1083*	0.0138	0.0857*	-0.0538*	-0.1835*	0.2568*	0.2863*	0.0551*	0.1646*	-0.0922*	0.0738*	-0.0796*	-0.0156	-0.2333*	-0.2488*	1		
phu yen	-0.1358*	-0.1653*	-0.0392*	-0.0702*	0.0259	-0.1868*	-0.0507*	-0.0720*	0.0476*	0.0236*	-0.0990*	-0.1911*	-0.0142	-0.0112	0.1553*	0.0489*	0.0746*	0.0250*	-0.2414*	-0.2445*	-0.2531*	1	
ben tre	0.0253*	-0.0433*	-0.0280*	0.0675*	0.0604*	-0.0132	0.0491*	0.00160	0.00250	-0.1869*	-0.1411*	-0.1121*	-0.1077*	-0.2827*	-0.2104*	-0.1115*	-0.1219*	0.0414*	-0.2484*	-0.2532*	-0.2622*	-0.2576*	1

Source: Author's calculation from the Young Lives in Vietnam

A.2. Factor analysis for the composite measure of the cognitive skill

(principal component factors; 1 factor retained)

Factor	Eigenvalue	Difference	Proportion	Cumulative
1	1.53568	1.07137	0.7678	0.7678
2	0.46432	.	0.2322	1.0000

Factor Loadings

Variable	1	Uniqueness
Math R2	0.87627	0.23216
PPVT R2	0.87627	0.23216

(principal component factors; 1 factor retained)

Factor	Eigenvalue	Difference	Proportion	Cumulative
1	1.57213	1.14426	0.7861	0.7861
2	0.42787	.	0.2139	1.0000

Factor Loadings

Variable	1	Uniqueness
Math R3	0.88660	0.21394
PPVT R3	0.88660	0.21394

A.3. Test for the significance of variables included in the model

Variable	p-value
Male	0.194
Minority	0.057
Father's education	0.000
Mother's education	0.000
Caregiver's aspiration for child's education (R2)	0.000
Math	0.000
PPVT	0.000
Self-efficacy	0.065
Self-esteem	0.038
Inclusion	0.949
Cognitive skill	0.000
Non-cognitive skill	0.075
Educational aspiration	0.000
Lagged wealth index	0.000
Lagged household size	0.002
Hours spent on household's chores	0.003
Hours spent on domestic tasks	0.006
Hours spent on paid activities	0.007
Urban	0.880
Lao Cai	0.000
Hung Yen	0.008
Phu Yen	0.078
Ben Tre	0.170

Source: Author's calculation from the Young Lives in Vietnam

A.4. Test for proportional hazard assumption based on Schoenfeld residuals

	Single skills				Composite skills			
	Without TVC		With TVC		Without TVC		With TVC	
	(S13)	(S14)	(S15)	(S16)	(S20)	(S21)	(S22)	(S23)
Math	0.866	0.975	0.522	0.638				
PPVT	0.300	0.136	0.204	0.0938				
Self-efficacy	0.593	0.662	0.510	0.550				
Self-esteem	0.534	0.601	0.654	0.766				
Inclusion	0.0335	0.0837	0.0350	0.0772				
Cognitive skill					0.342	0.222	0.432	0.272
Non-cognitive skill					0.235	0.353	0.207	0.325
Educational aspiration	0.601	0.360	0.470	0.288	0.806	0.548	0.688	0.462
Male	0.0969	0.151	0.0684	0.106	0.0523	0.110	0.0362	0.0802
Minority	0.255	0.744	0.261	0.739	0.332	0.827	0.338	0.825
Father's education	0.285	0.222	0.308	0.215	0.294	0.191	0.309	0.190
Mother's education	0.104	0.150	0.104	0.135	0.102	0.110	0.104	0.108
Caregiver's aspiration	0.769	0.709	0.799	0.769	0.618	0.587	0.609	0.584
Prior wealth index	0.785	0.416	0.969	0.323	0.737	0.473	0.828	0.406
Prior household size	0.268	0.756	0.0786	0.339	0.276	0.769	0.101	0.402
Household's chores	0.804	0.627	0.844	0.693	0.697	0.956	0.683	0.916
Domestic tasks	0.523	0.508	0.637	0.755	0.681	0.662	0.751	0.826
Paid activities	0.578	0.854	0.478	0.714	0.496	0.829	0.409	0.719
Urban areas	0.764		0.793		0.810		0.853	
Lao Cai		0.616		0.548		0.510		0.500
Hung Yen		0.433		0.440		0.486		0.471
Phu Yen		0.867		0.824		0.947		0.946
Ben Tre		0.107		0.131		0.124		0.148
Global test	0.556	0.523	0.339	0.426	0.635	0.575	0.471	0.523

Note: Numbers reported in the cells are the Chi squared p-values of the variables included in the models; The number part in the label of each column refers to the model for which the test is conducted, e.g. S13 refers to the result of test for Model (13)

Source: Author's calculation from the Young Lives in Vietnam

A.5. Test for proportional hazard assumption for Models (14) and (21)

	Single skills (13)			Composite skills (19)		
	Without TVC	With TVC		Without TVC	With TVC	
Math	-0.5418*** (0.1738)	-0.8146 (1.3344)	0.0307 (0.1656)			
PPVT	0.0077 (0.1560)	1.6111 (1.2346)	-0.1947 (0.1465)			
Self-efficacy	-0.0546 (0.2530)	0.1519 (1.9233)	-0.0252 (0.2382)			
Self-esteem	-0.2918 (0.2416)	-1.0412 (1.8353)	0.0884 (0.2216)			
Inclusion	0.3248 (0.2567)	3.8101* (2.0664)	-0.4154* (0.2428)			
Cognitive skill				-0.3657** (0.1597)	1.2688 (1.3419)	-0.2114 (0.1702)
Non-cognitive skill				0.0202 (0.3177)	1.9112 (2.4063)	-0.2361 (0.2936)
Educational aspiration	-0.6235* (0.3467)	-2.7687 (2.4898)	0.2592 (0.3054)	-0.5856 (0.3569)	-2.5110 (2.4590)	0.2382 (0.2997)
Male	0.0058 (0.2930)	-0.0508 (0.2956)		0.0613 (0.2902)	0.0396 (0.2939)	
Minority	-0.6921 (0.5540)	-0.8471 (0.6049)		-0.7467 (0.5709)	-0.8376 (0.6017)	
Father's education	-0.4436** (0.2078)	-0.4358** (0.2107)		-0.4510** (0.2068)	-0.4611** (0.2101)	
Mother's education	-0.3009 (0.2432)	-0.3059 (0.2453)		-0.3062 (0.2386)	-0.2805 (0.2407)	
Caregiver's aspiration	-0.3362 (0.3523)	-0.2859 (0.3667)		-0.4802 (0.3562)	-0.4714 (0.3642)	
Prior wealth index	-1.7510 (1.1319)	2.7858 (6.3347)	-0.5639 (0.7912)	-1.6578 (1.1296)	2.8171 (6.1920)	-0.5507 (0.7695)
Prior household size	0.1530* (0.0852)	0.8477 (0.5961)	-0.0874 (0.0767)	0.1447* (0.0860)	0.8255 (0.5834)	-0.0864 (0.0749)
Household's chores	-0.3170* (0.1662)	-0.8610 (1.2035)	0.0604 (0.1446)	-0.2526 (0.1584)	-0.2479 (1.1407)	-0.0027 (0.1381)

Domestic tasks	0.1589** (0.0766)	-0.1218 (0.6719)	0.0275 (0.0752)	0.1824** (0.0759)	0.1805 (0.6556)	-0.0025 (0.0733)
Paid activities	0.0662 (0.0745)	1.9801 (1.8455)	-0.2130 (0.2055)	0.0783 (0.0728)	2.7028 (1.9212)	-0.2917 (0.2142)
Lao Cai	-2.0899*** (0.7365)	-1.9509** (0.7738)		-2.2369*** (0.7446)	-2.1308*** (0.7695)	
Hung Yen	-1.7080** (0.8033)	-1.5984** (0.8150)		-1.7818** (0.7990)	-1.7466** (0.8079)	
Phu Yen	-0.9545** (0.4812)	-0.7703 (0.5002)		-0.9336** (0.4751)	-0.8130* (0.4872)	
Ben Tre	-0.9048 (0.5529)	-0.7975 (0.5773)		-1.0050* (0.5517)	-0.8896 (0.5667)	
No. of obs.	4,854	4,854		4,854	4,854	
No. of failures	53	53		53	53	
No. of subjects	780	780		780	780	
Log-Likelihood	-286.25	-281.74		-290.25	-287.24	
Chi2 p-value	0.0000	0.0000		0.0000	0.0000	
Adjusted R2	0.174	0.187		0.163	0.171	

Note: *** p<0.01, ** p<0.05, * p<0.1

Source: Author's calculation from the Young Lives in Vietnam

A.6. Robustness check – Cox models with single measures of cognitive skill

	(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)	(viii)	(ix)	(x)	(xi)
									TVC		TVC
Math	-0.7097*** (0.1564)		-0.6817*** (0.1565)	-0.5523*** (0.1637)	-0.4667*** (0.1636)	-0.5057*** (0.1652)	-0.5755*** (0.1736)	-0.5775*** (0.1735)	-0.0708*** (0.0217)	-0.5418*** (0.1738)	-0.0669*** (0.0217)
PPVT	-0.3083** (0.1387)		-0.2951** (0.1393)	-0.2238 (0.1510)	-0.0874 (0.1525)	-0.0298 (0.1471)	0.0349 (0.1508)	0.0141 (0.1529)	0.0026 (0.0184)	0.0077 (0.1560)	0.0011 (0.0187)
Self-efficacy		-0.3721 (0.2377)	-0.2637 (0.2406)	-0.1256 (0.2368)	-0.0840 (0.2451)	-0.1180 (0.2385)	-0.0316 (0.2476)	-0.0579 (0.2517)	-0.0087 (0.0317)	-0.0546 (0.2530)	-0.0069 (0.0317)
Self-esteem		-0.5133** (0.2340)	-0.3345 (0.2430)	-0.3347 (0.2474)	-0.2551 (0.2410)	-0.2381 (0.2422)	-0.3185 (0.2437)	-0.3068 (0.2448)	-0.0354 (0.0300)	-0.2918 (0.2416)	-0.0333 (0.0296)
Inclusion		0.3028 (0.2486)	0.2496 (0.2477)	0.3155 (0.2418)	0.3477 (0.2422)	0.3194 (0.2429)	0.3642 (0.2431)	0.4127* (0.2507)	0.0406 (0.0301)	0.3248 (0.2567)	0.0303 (0.0308)
Educational aspiration				-1.1199*** (0.3235)	-0.6711** (0.3378)	-0.6561* (0.3414)	-0.6669** (0.3400)	-0.6521* (0.3417)	-0.0774* (0.0426)	-0.6235* (0.3467)	-0.0732* (0.0434)
Male				0.2868 (0.2824)	0.1860 (0.2823)	0.2110 (0.2826)	0.0680 (0.2906)	0.0245 (0.2914)	0.0437 (0.2915)	0.0058 (0.2930)	0.0164 (0.2932)
Minority				-0.2744 (0.4021)	-0.6713* (0.3995)	-1.1260** (0.4511)	-1.1372** (0.4794)	-1.1759** (0.4932)	-1.1498** (0.5004)	-0.6921 (0.5540)	-0.6610 (0.5551)
Father's education					-0.3666* (0.2048)	-0.3555* (0.2062)	-0.3603* (0.2065)	-0.3879* (0.2061)	-0.3793* (0.2057)	-0.4436** (0.2078)	-0.4429** (0.2073)
Mother's education					-0.5598** (0.2291)	-0.4651** (0.2370)	-0.4721** (0.2371)	-0.4459* (0.2343)	-0.4565* (0.2339)	-0.3009 (0.2432)	-0.2941 (0.2433)
Caregiver's aspiration					-0.4000 (0.3404)	-0.3140 (0.3483)	-0.3258 (0.3472)	-0.3457 (0.3509)	-0.3291 (0.3547)	-0.3362 (0.3523)	-0.3161 (0.3569)
Prior wealth index						-1.1727	-0.7988	-1.7527*	-0.2301*	-1.7510	-0.2370*

						(0.9287)	(0.9445)	(1.0287)	(0.1290)	(1.1319)	(0.1419)
Prior household size						0.2182**	0.1907**	0.1764**	0.0204*	0.1530*	0.0187*
						(0.0855)	(0.0855)	(0.0854)	(0.0113)	(0.0852)	(0.0112)
Household's chores							-0.3681**	-0.3431**	-0.0391**	-0.3170*	-0.0370*
							(0.1667)	(0.1630)	(0.0198)	(0.1662)	(0.0202)
Domestic tasks							0.0864	0.1066	0.0096	0.1589**	0.0159*
							(0.0740)	(0.0749)	(0.0087)	(0.0766)	(0.0087)
Paid activities							0.0534	0.0600	0.0053	0.0662	0.0059
							(0.0726)	(0.0730)	(0.0082)	(0.0745)	(0.0084)
Urban areas								1.0343**	0.9915**		
								(0.4371)	(0.4301)		
Lao Cai										-2.0899***	-2.1514***
										(0.7365)	(0.7396)
Hung Yen										-1.7080**	-1.6980**
										(0.8033)	(0.8007)
Phu Yen										-0.9545**	-0.8877*
										(0.4812)	(0.4719)
Ben Tre										-0.9048	-0.8780
No. of obs.	4,854	4,854	4,854	4,854	4,854	4,854	4,854	4,854	4,854	4,854	4,854
No. of failures	53.0000	53.0000	53.0000	53.0000	53.0000	53.0000	53.0000	53.0000	53.0000	53.0000	53.0000
No. of subjects	780.0000	780.0000	780.0000	780.0000	780.0000	780.0000	780.0000	780.0000	780.0000	780.0000	780.0000
Log-Likelihood	-320.38	-342.37	-318.67	-311.37	-299.39	-295.85	-291.77	-289.23	-290.51	-286.25	-287.00
Chi2 p-value	0.0000	0.0375	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Adjusted R2	0.076	0.012	0.081	0.102	0.136	0.146	0.158	0.166	0.162	0.174	0.172

Note: *** p<0.01, ** p<0.05, * p<0.1

Source: Author's calculation from the Young Lives in Vietnam

A.7. Robustness check – Cox models with composite measures of cognitive skill

	(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)	(viii)	(ix)	(x)	(xi)
									TVC		TVC
Cognitive skill	-0.7920*** (0.1079)		-0.7778*** (0.1093)	-0.5842*** (0.1356)	-0.3854*** (0.1431)	-0.3512** (0.1489)	-0.3622** (0.1538)	-0.4048** (0.1580)	-0.0542*** (0.0200)	-0.3657** (0.1597)	-0.0497** (0.0202)
Non-cognitive skill		-0.5859* (0.3021)	-0.3433 (0.3143)	-0.1148 (0.3182)	0.0150 (0.3115)	-0.0087 (0.3079)	0.0647 (0.3102)	0.0956 (0.3187)	0.0057 (0.0391)	0.0202 (0.3177)	-0.0013 (0.0388)
Educational aspiration				-1.2043*** (0.3236)	-0.6739** (0.3422)	-0.6538* (0.3478)	-0.6352* (0.3505)	-0.6015* (0.3534)	-0.0713 (0.0437)	-0.5856 (0.3569)	-0.0680 (0.0443)
Male				0.2715 (0.2801)	0.2403 (0.2794)	0.2929 (0.2799)	0.1640 (0.2864)	0.1206 (0.2869)	0.1273 (0.2871)	0.0613 (0.2902)	0.0648 (0.2904)
Minority				-0.1718 (0.3941)	-0.6220 (0.3920)	-1.1073** (0.4595)	-1.2258** (0.4949)	-1.3043** (0.5130)	-1.3068** (0.5223)	-0.7467 (0.5709)	-0.7568 (0.5728)
Father's education					-0.3508* (0.2028)	-0.3489* (0.2053)	-0.3553* (0.2055)	-0.3793* (0.2047)	-0.3710* (0.2046)	-0.4510** (0.2068)	-0.4452** (0.2062)
Mother's education					-0.5683** (0.2256)	-0.4847** (0.2338)	-0.4908** (0.2325)	-0.4540** (0.2311)	-0.4580** (0.2309)	-0.3062 (0.2386)	-0.3005 (0.2387)
Caregiver's aspiration					-0.5150 (0.3400)	-0.4546 (0.3494)	-0.4580 (0.3521)	-0.4797 (0.3561)	-0.4505 (0.3583)	-0.4802 (0.3562)	-0.4546 (0.3587)
Prior wealth index						-0.9282 (0.9295)	-0.5451 (0.9552)	-1.4992 (1.0349)	-0.1982 (0.1284)	-1.6578 (1.1296)	-0.2262 (0.1403)
Prior household size						0.2125** (0.0851)	0.1932** (0.0859)	0.1774** (0.0860)	0.0207* (0.0112)	0.1447* (0.0860)	0.0177 (0.0112)
Household's chores							-0.2795* (0.1585)	-0.2602* (0.1552)	-0.0307 (0.0189)	-0.2526 (0.1584)	-0.0307 (0.0192)
Domestic tasks							0.1079	0.1306*	0.0115	0.1824**	0.0180**

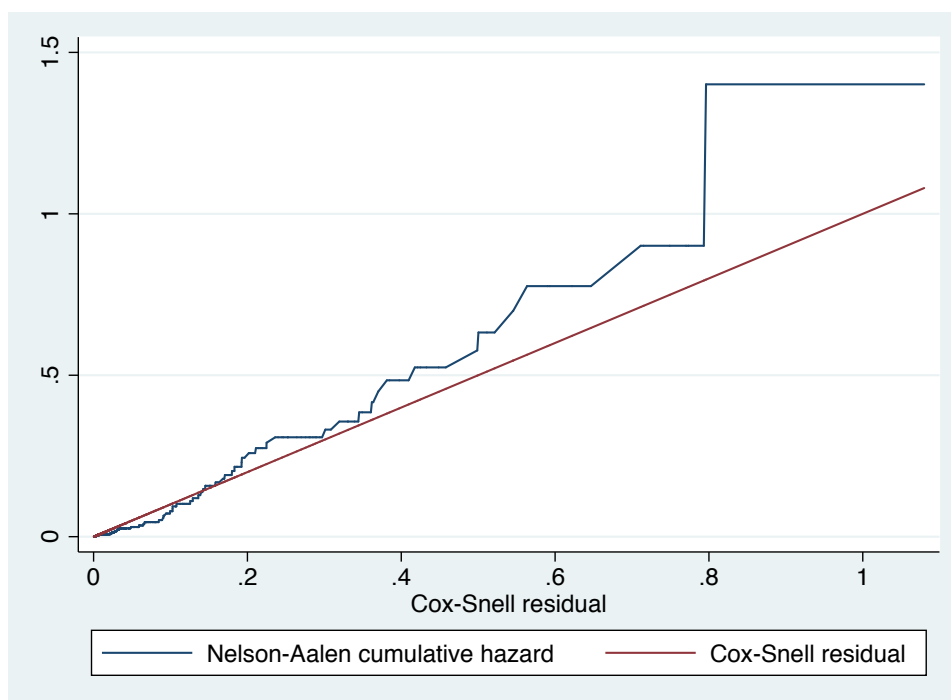
							(0.0720)	(0.0733)	(0.0084)	(0.0759)	(0.0086)
Paid activities							0.0717	0.0776	0.0069	0.0783	0.0069
							(0.0707)	(0.0713)	(0.0080)	(0.0728)	(0.0082)
Urban areas								1.0605**	1.0308**		
								(0.4369)	(0.4295)		
Lao Cai										-2.2369***	-2.2980***
										(0.7446)	(0.7543)
Hung Yen										-1.7818**	-1.7642**
										(0.7990)	(0.7972)
Phu Yen										-0.9336**	-0.8958*
										(0.4751)	(0.4666)
Ben Tre										-1.0050*	-0.9986*
No. of obs.	4,854	4,854	4,854	4,854	4,854	4,854	4,854	4,854	4,854	4,854	4,854
No. of failures	53.0000	53.0000	53.0000	53.0000	53.0000	53.0000	53.0000	53.0000	53.0000	53.0000	53.0000
No. of subjects	780.0000	780.0000	780.0000	780.0000	780.0000	780.0000	780.0000	780.0000	780.0000	780.0000	780.0000
Log-Likelihood	-325.02	-344.77	-324.43	-316.04	-303.12	-299.97	-296.47	-293.81	-294.17	-290.25	-290.27
Chi2 p-value	0.0000	0.0554	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Adjusted R2	0.062	0.005	0.064	0.088	0.125	0.135	0.145	0.152	0.151	0.163	0.163

Note: *** p<0.01, ** p<0.05, * p<0.1

Source: Author's calculation from the Young Lives in Vietnam

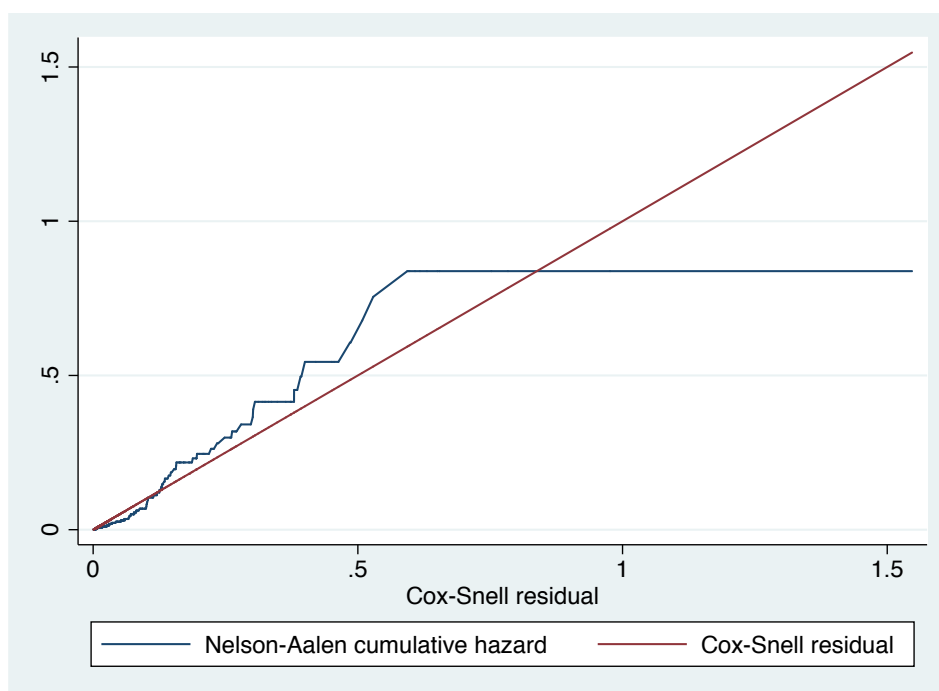
B. Figures

B.1. Goodness of fit of the full model (7) – individual measures of skills



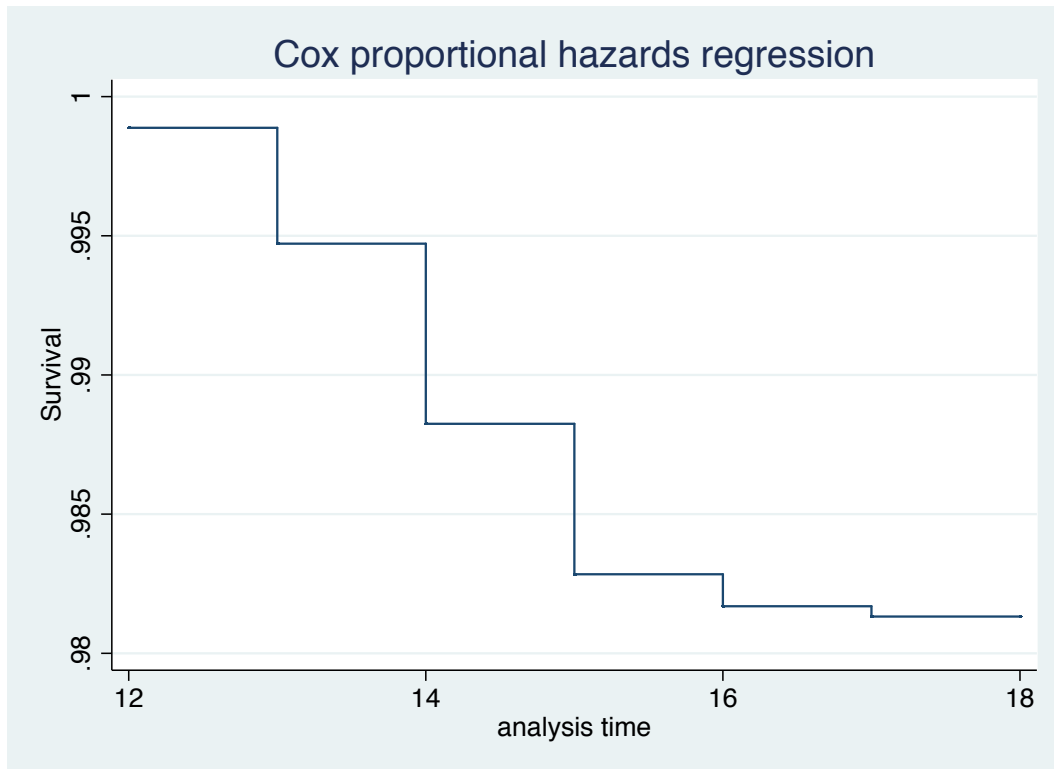
Source: Author's calculation from the Young Lives in Vietnam

B.2. Goodness of fit of the full model (7)– composite measures of skills



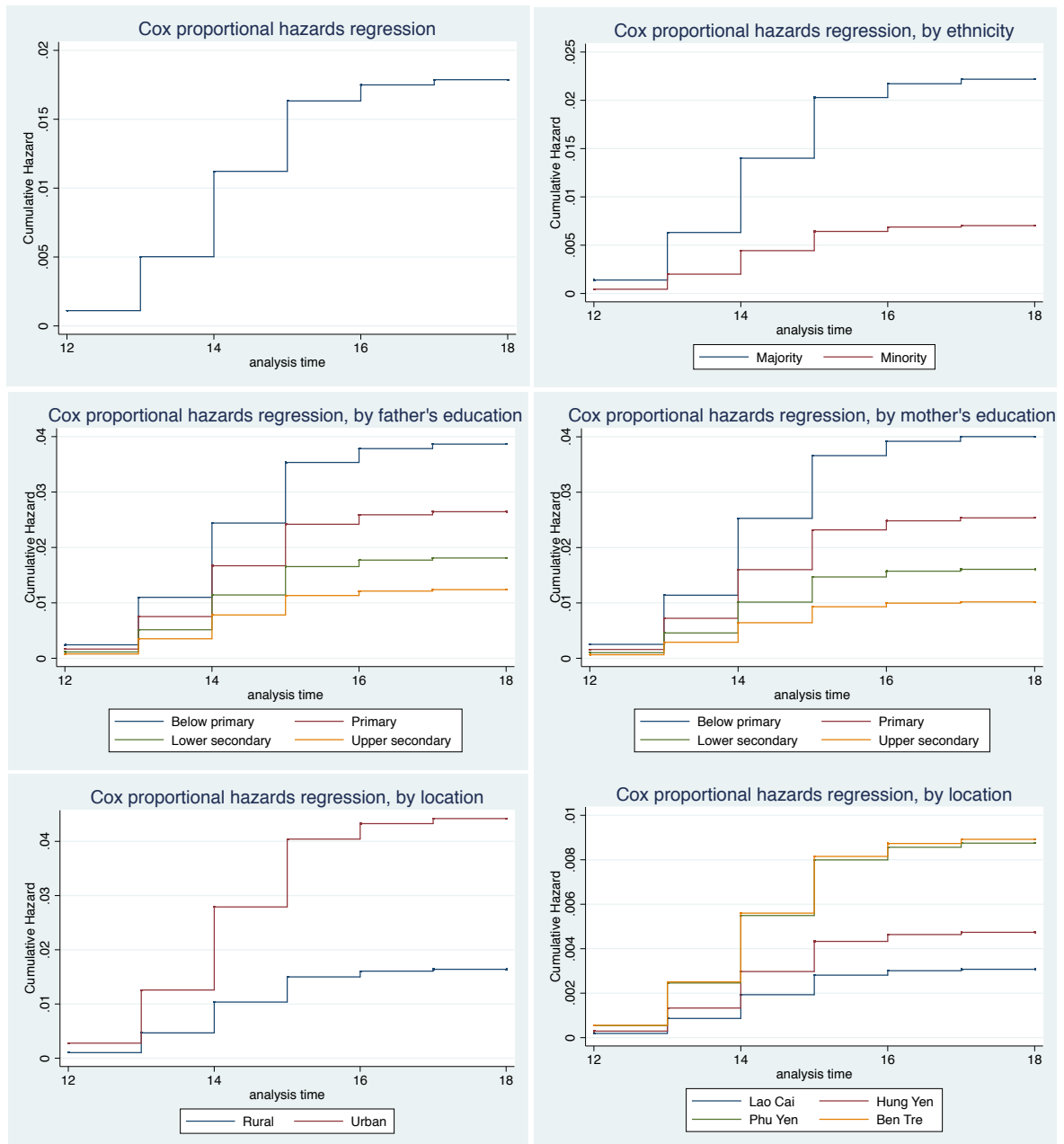
Source: Author's calculation from the Young Lives in Vietnam

B.3. Estimated survival function



Source: Author's calculation from the Young Lives in Vietnam

B.4. Estimated cumulative hazard, whole sample and subgroups



Source: Author's calculation from the Young Lives in Vietnam

Chapter 2: Does full-day schooling reduce educational inequality in Vietnam?

Abstract

Full-day schooling was introduced in Vietnam to deal with the lack of instructional time in primary education compared to international standards. Yet full-day schooling could impact educational inequality by filling the gap in instructional time between children from different family backgrounds, given that well-off families tend to offset the lack of instructional time with private tutoring. This paper draws on data from the Young Lives School Survey 2011-2012 to investigate whether the massive rollout of full-day schooling in recent years can improve equality of opportunity in Vietnam. First, it examines the variation in instructional time and school resources between pupils with and without full-day schooling from different social backgrounds. It then investigates the learning progress associated with full-day schooling across social backgrounds. The evidence suggests that full-day schooling has not narrowed the inequality of opportunity in learning achievement at this stage of its implementation. It may, in fact, have the opposite effect by actually broadening the gap in learning progress between children from different social backgrounds.

1. Introduction

Many developing countries are forced by financial and human resources shortages to restrict instructional time to half a day. Yet as these countries develop and attain universal enrolment in primary education, many of them introduce policies to lengthen instructional time in order to align with international standards of education. Vietnam is one such country. Although the full-day schooling policy was decided on very early in Vietnam (1956), its implementation remained marginal until the 2000s when it was scaled up by the School Education Quality Assurance Program starting in 2009.

Full-day schooling is expected to improve children's learning achievements by lengthening instructional time. This is borne out by most of the studies on the subject, although the observed effect of longer instructional time on achievement is often small (see Holland *et al.*, 2015 and Orkin, 2013, for a review). Evidence on Vietnam finds a close association between full-day schooling and achievement (Tran, 2014; World Bank, 2011), yet none of the studies addresses causal inference.

Lengthening instructional time has also been proposed as a policy option to reduce inequality of opportunity, especially in developed countries like the USA (Biernat, 2011). Inequality of opportunity is defined here as the association between educational performance and family background. The argument put forward is that this policy narrows the gap in instructional time between children from disadvantaged backgrounds and well-off children. Well-off families can offset a lack of instructional time with private tutoring and extra classes. Poor families, however, rely mostly on schools to provide academic education to their children, unlike children from well-off and middle-class families who rely on school for just a portion of their learning and are more likely to receive intellectual stimulation during the time they spend at home. Replacing home time with school time is supposed to give children from different social backgrounds the benefit of a more similar learning environment. However, narrowing the gap in instructional time across children from different social backgrounds will reduce inequality of opportunity in learning achievement only if the quality of the instructional time is the same for all groups of children.

Very few studies have analysed the effect of lengthening instructional time on inequality of opportunity, particularly in developing countries, and the studies that have done so reach different conclusions. Schutz *et al.* (2008) look at the association between social background and performance in mathematics in 54 countries and find no significant difference in equality of opportunity between half-day and whole-day school systems. Their study is based on comparable data from the Third International Mathematics and Science Study (TIMSS). Yet the study presents some limitations due to data constraints. In particular, the measurement of social background in the form of the number of books at home is simplistic. In addition, the fact that the authors use country fixed effects means that they are able to control for unobserved country heterogeneity only when it does not vary over time.

Hindicapie (2016) shows that switching from a half school day to a full school day in Colombia improves the lowest-income pupils' achievements more and therefore reduces inequality of opportunity in learning achievement. Their demonstration relies on an identification strategy based on within-school variation in school day length, which is assumed to be exogenous. Lavy (2010) takes PISA 2006 data from over 50 countries and also finds that instructional time has a greater effect on

the academic achievements of pupils from low socioeconomic status families. The effect of instructional time is identified based on a comparison of the performance of the same pupils in different subjects in order to overcome potential selection and endogeneity problems. Lavy (2012) comes to the same conclusion for Israel using another identification strategy based on a natural experiment providing the opportunity to exploit a sharp change in the method of funding schools in Israel.

Lastly, Orkin (2013) studies the effect of a massive switch to full-day schooling in a developing country: Ethiopia. Causal inference is drawn using a natural experiment in the form of the variation in reform implementation time by district. This variation is assumed to be exogenous. The author's result differs from other studies in that she concludes there is an increase in inequality of opportunity: the effects of full-day schooling are greater among better-off children, defined as children who are not stunted, children from richer households and children in urban schools.

The purpose of this paper is to contribute to this question by addressing the case of Vietnam, a developing country with a particularly interesting context. First, inequality of opportunity in education is a growing concern for the country. Inequality in education across socioeconomic groups and across ethnic groups has increased in recent decades, especially in secondary education attainment (World Bank, 2011). Second, the introduction of full-day schooling in Vietnam presents a particularity that may change the effect of full-day schooling on inequality of opportunity. Family participation in funding the full-day schooling policy is far from negligible. This could inhibit the policy's positive impact on equality of opportunity. On the other hand, Vietnam conducted an ambitious policy from 2009 to 2015 to support schools in the most disadvantaged provinces in their switch to full-day schooling.

This paper therefore examines whether full-day schooling is associated with greater equality of opportunity in educational achievement. Full-day schooling reduces inequality of opportunity in learning achievement only if pupils from disadvantaged backgrounds benefit more from this policy than their well-off counterparts in terms of learning time and resources at school. To test this, we inspect the variations in instructional time and school resources between pupils with and without full-day schooling from different social backgrounds. The data are taken from the 2011 School Survey collected by the Young Lives Project in Vietnam. The learning progress associated with full-day schooling across social backgrounds is investigated by estimating value-added models by subgroup of social backgrounds. Pupils' past educational inputs and unobserved characteristics are partially captured by lagged achievement, and biases due to the selection process are mitigated by 2SLS estimation. However, unobserved time-varying characteristics are not controlled for. Thus, this paper provides an array of evidence on the relationship between full-day schooling and inequality in learning progress across levels of social backgrounds, rather than causal inferences.

The results show that full-day schooling only partially narrows the gap in instructional time among students from different social backgrounds. Pupils from well-off families attend both full-day schooling and extra classes, while the poorest children still have less access to full-day schooling. In addition, although the implementation of full-day schooling is associated with more school and class resources, resources are unequally distributed across full-day schooled students from different social backgrounds. The higher the students' social background, the greater the resources to which they

have access. Lastly, full-day schooling is not associated with better learning progress, especially among children from disadvantaged backgrounds. This evidence illustrates the minimal contribution of FDS to reducing inequality of opportunity in learning achievement. It even appears to contribute to this inequality.

The paper is organized as follows. Section 2 provides an overview of the introduction of full-day schooling in Vietnam. The data and definitions used in the analysis are described in Section 3. Section 4 examines whether full-day schooling reduces instructional time among students from different social backgrounds. The distribution of additional resources allocated for full-day schooling is analysed across social backgrounds in section 5. Section 6 considers whether the effect of full-day schooling on learning achievement varies by the children's social background. Section 7 summarizes the results and presents their policy implications.

2. Full-day schooling in the context of Vietnam

Instructional time in Vietnamese schools lag far behind both the advanced economies and other Southeast Asian countries (Appendix B.2 and B.3). The time constraint prevents teachers from covering the basic primary curriculum. Concurrent private tutoring is widespread, with the share of primary school pupils attending extra classes reaching 35 percent in 2012,³⁶ increasing the well-off children's instructional time. Aware of these disparities, the government has been promoting a transition to full-day schooling with a minimum national standard of 35 periods per week by 2025, with each period including 35 to 40 minutes of instructional time (Vietnam Education Development Strategic Plan 2008-2020). This implies an increase of at least 40 percent in instructional time to make the switch from half-day schooling to full-day schooling at the primary level.

Following its authorization in 1956,³⁷ full-day schooling in primary education was introduced on a voluntary basis in Vietnam mostly in urban areas through to the 2000s. The adoption of full-day schooling is decided at the school level with limited support from the government. Local authorities and parents share the cost of the infrastructure and additional teacher remuneration induced by full-day schooling. Full-day schooling is introduced either for the entire school or some classes in the school, and the lengthening of learning time can vary from 30 to 40 periods per week depending on the school's volition and resources. Consequently, the first schools to lengthen learning time were those where parents were willing and able to pay the extra costs. The proportion of pupils attending full-day schooling in poor districts remained very low until 2009, standing at 30 percent in 2008 compared to 59 percent nationwide (World Bank, 2011).

³⁶ Authors' estimations from the 2012 Vietnam Household Living Standards Survey.

³⁷ Ministry of Education Decree No. 596-NĐ of 30 August 1956 on issuing the regulation for ten-year general education authorised full-day schooling for students in schools with enough classrooms, with no further guidance on the implementation of full-day schooling.

In late 2009, the Government launched the School Education Quality Assurance Program³⁸ (SEQAP) to improve the quality of primary education by supporting the transition to full-day schooling in primary schools. The program lasted for six years (ending in 2015) and was implemented in the country's 36 most disadvantaged of its 63 provinces. Some 1,770 schools benefited from this program. The total cost of SEQAP was USD 186 million. SEQAP subsidized the training and additional working hours of teachers and school principals in all schools switching to full-day schooling. Funding concentrated on those schools with just 23 periods of instructional time per week that chose to switch to full-day schooling, especially in ethnic minority communities. The Ministry of Education and Training reported that the number of schools switching to full-day schooling grew quickly, with 73 percent of pupils attending full-day schooling in primary school in 2012, and 70 percent of pupils in this case in the poorest districts (Bodewig *et al.*, 2014). The massive extension of full-day schooling in recent years, especially for the most disadvantaged children, calls for an investigation into its relationship with educational inequality.

3. Data and definitions

The data source used in this paper is the School Survey conducted in Vietnam in the 2011-2012 school year. The School Survey's sampling is based on one of the Young Lives longitudinal surveys. Any school attended by at least one younger cohort child in Grade 5 in 2011 was included in the school survey sample. The sample of children surveyed was then expanded by adding classmates randomly selected to conduct school and class level analyses. In total, 20 children per class are surveyed. The School Survey sample consists of 3,284 Grade 5 pupils in 176 classes at 52 schools. Among these pupils, 1,125 are part of the Young Lives longitudinal survey sample.

The Young Lives School Survey 2011 is a highly suitable source of data to investigate the effect of full-day schooling on inequality of opportunity as it provides rich information on instructional time and the characteristics of schools, principals, teachers and students. In addition to students, the survey also interviews school principals and head teachers.³⁹ More importantly, students' academic achievements recorded at two points in time means that their academic progress can be analysed.

The data are used for the analysis based on the core definitions as follows:

³⁸ A multi-donor grant-funded program co-financed and co-implemented by the Government of Vietnam and its partners (World Bank, DFID and Belgium).

³⁹ Head teachers in Vietnam are the main teachers who teach the core subjects of math, Vietnamese, history, etc. Other teachers provide a support role, teaching music, art, etc. Head teachers are also responsible for managing the class and for communicating with parents.

Full-day schooling (FDS)

Looking at the data, the availability of FDS to a student can be determined in four ways. The school principal, the head teacher and the student him/herself are asked directly whether FDS is available at the school level, at the class level or for the student respectively. The fourth way to define FDS is to use the number of periods that students have at school in one week as reported by the head teacher.⁴⁰ We choose this last variable for several reasons. First, FDS may have been implemented for a whole school, or some selected classes, or even for some students in one class. Thus, the principal's positive answer as to the availability of FDS in the school may conceal different situations across classes and does not guarantee that the student in question benefits from FDS. To a lesser extent, the same applies to the teacher's answer. The student's answer to the question "Do you attend full-day schooling (two shifts - morning and afternoon)?" can be biased if the student confuses FDS with extra classes given by the teacher in the classroom. In addition, these three variables based on direct statements do not provide a definition which guarantees that students with FDS have longer instructional time than students without FDS. Thus, FDS is defined based on the class timetable reported by the head teacher: a student is considered as attending FDS if the head teacher declares that the student is taught at least 30 periods per week. The threshold of 30 periods has been retained firstly because it is the minimum government requirement to qualify as full-day schooling. Second, it is the most consistent with the common declaration of the head teacher and the principal.⁴¹ Based on this definition, 40.5 percent of students receive FDS.⁴²

Extra classes

Extra classes refer to classes that students take in addition to their formal schooling day. Responses by principals in the survey suggest that less than 4 percent of the schools arrange non-compulsory additional classes by the same teacher who teaches during the normal schooling day. Head teachers only gave information on their private tuition in general rather than for the students in

⁴⁰ For each day of the week, the head teacher was asked "*How many periods does this grade 5 class have on each of these days?*"

⁴¹ According to this alternative criterion, FDS is defined when the head teacher's answer to the question "*Do the pupils in your class attend full-day schooling?*" is "*Yes, all pupils attend full-day schooling*" and the principal's answer to the question "*Does your school offer full-day schooling to any pupils in Grade 5?*" is "*Yes, all pupils*". When the two criteria are compared, the exclusion errors for the criteria based on 30 periods cover 14% of the students and the inclusion errors just 5%. When the threshold is increased to 35 periods (the government target), exclusion errors cover 25% of the students.

⁴² This proportion differs from the 73% of pupils attending full-day schooling in primary schools reported by the Ministry of Education and Training for three main reasons: (i) the School Survey sample is pro-poor biased as mentioned in the description of the data; (ii) only Grade 5 pupils are covered by the School Survey rather than all primary school pupils as in the MOET figures; (iii) the definition of full-day schooling may be different in the two sources of data because of the broad variations in FDS practice and the unclear distinction between formal full-day schooling and informal extra classes.

the class. Hence this study defines extra classes based on the pupils' statement as to whether they attend extracurricular/non-compulsory extra classes.⁴³ A total of 39 percent of the students attend extra classes.

Learning achievement

Student academic achievement is measured by test scores in mathematics and Vietnamese recorded at two points in time, one at the beginning and the other at the end of the 2011-2012 school year. Each test is a 45-minute test containing 30 multiple-choice items. These tests were designed by the Vietnamese National Institute of Educational Sciences to cover key subject areas in the Grade 5 curriculum. The first tests cover the Grade 4 curriculum to measure learning level at Grade 5 entry. The second tests reflect the G5 curriculum, more advanced Grade 4 questions and common items replicated from the first tests. There are 15 and 12 common items in the first and second tests in Mathematics and in Vietnamese respectively.⁴⁴ These are used to transform the first and second test scores into a common interval scale using IRT.⁴⁵ The two rounds of tests track the students' academic achievements over the year.

Instructional time

The time of instruction that a pupil receives from his/her attendance at formal schooling and extra classes constitutes his/her instructional time. In this study, instructional time is measured by the number of periods per week in schooling and in extra classes.

Social background

The pupils' social background⁴⁶ is measured by a composite index to approximate their socioeconomic characteristics. Social background level is constructed using survey information collected on the student's home and family, including the parents' educational level, number of meals

⁴³ For information on extra classes, the principals are asked *"Does your school offer non-compulsory additional classes available outside of the normal school day, provided at an additional charge to any students who wish to participate, on the premises?"* and *"If yes, are these additional classes taught by the same teachers who teach during the normal school day?"*. Teachers are just asked *"Do you do private tuition to supplement your income?"* For students, the question is *"Do you attend extracurricular/non-compulsory extra classes, whether at school or not?"* and *"How many hours do you attend classes each week in each of these subjects – Math, Vietnamese and Other subjects?"*

⁴⁴ See (Rolleston *et al.*, 2013) for a detailed description of the tests.

⁴⁵ Item Response Theory is "the model-based measurement in which trait level estimates depend on both a person's responses and the properties of the items that were administered" (Embretson & Reise, 2000). Documentation explaining the development and validation of the IRT in the Vietnam School Survey can be downloaded from:

http://www.younglives.org.uk/sites/www.younglives.org.uk/files/Vietnam-School-Survey_R1_Note-on-Item-Response-Scoring.pdf

⁴⁶ Also referred to as family background or socioeconomic status in relevant empirical literature.

per day, asset-based factors and factors relating to the student's learning environment at home (see detailed description in Appendix A.1). Given that these factors are categorical variables, the social background index is computed applying Multiple Correspondence Analysis (MCA).

MCA shows that 78 percent of the variation in social background factors is explained by the first dimension (Appendix A.2), providing grounds for using the first dimension alone to rank students' social background levels. Among the most important contributory factors, having a computer, Internet, an air conditioner at home and a mother or father with upper secondary education all contribute positively to the index while having no study desk, no study chair, no study lamp and no place to study contribute negatively (Appendix A.3). In other words, the higher the index, the higher the student's socioeconomic status. The average social background indices for students with these factors are presented in Table 2- 1. The distribution of students across the social background index is provided in Appendix B.1.

Table 2- 1: Average social background indices for students with selected factors

Characteristics	No	Yes	
A computer at home	-0.3658	1.1989	***
Internet at home	-0.3049	1.3196	***
Air conditioner at home	-0.1902	1.4459	***
Mother completed upper secondary education	-0.2735	0.8700	***
Father completed upper secondary education	-0.2939	0.7837	***
A study desk at home	-1.6068	0.2188	***
A study chair at home	-1.5108	0.2112	***
A study lamp at home	-0.9593	0.3618	***
Own place to study at home	-1.0526	0.2641	***

Note: *** Difference significant at the 1% level

Source: Authors' calculations based on the Young Lives School Survey 2011

4. Does FDS reduce the gap in instructional time between students with different social backgrounds?

The introduction of a longer school day by FDS could increase instructional time for children from low social backgrounds while keeping the time constant for high social background children for whom full-day schooling may replace their extra classes. Calculations from the data show that instructional time is still low in Vietnam and that the government target of 35 periods per week has been achieved only for the most advantaged quintile of students and when extra classes are taken into account (Table 2- 2). In addition, instructional time increases with social background level and the gap in instructional time across social backgrounds is still huge. On average, students in the first quintile of social

background have just 29 periods per week, while the figure for the highest social background quintile is 39.

Table 2- 2: Instructional time and access to FDS and extra classes by quintile of social background

Quintile of social background	Q1	Q2	Q3	Q4	Q5	All
Instructional time (number of periods of schooling and extra classes per week)	29.22	32.29	33.00	33.96	38.86	33.47
Share of students with FDS (%)	24.16	38.36	44.00	48.52	47.68	40.53
Share of students without FDS (%)	75.84	61.64	56.00	51.48	52.32	59.47
Total	100	100	100	100	100	100
Share of students with extra classes (%)	20.64	33.90	36.38	39.29	63.34	38.69
Share of students without extra classes (%)	79.36	66.10	63.62	60.71	36.66	61.31
Total	100	100	100	100	100	100
<i>Obs</i>	654	652	657	649	652	3264

Source: Authors' calculations based on the Young Lives School Survey 2011

Table 2- 3: Access to SEQAP schools by social background (percentage)

Quintile of social background	Non-SEQAP school	SEQAP School
Q1	19.92	20.80
Q2	19.64	22.22
Q3	19.22	26.24
Q4	19.78	20.57
Q5	21.44	10.17
Total	100	100
<i>Obs</i>	2 841	423

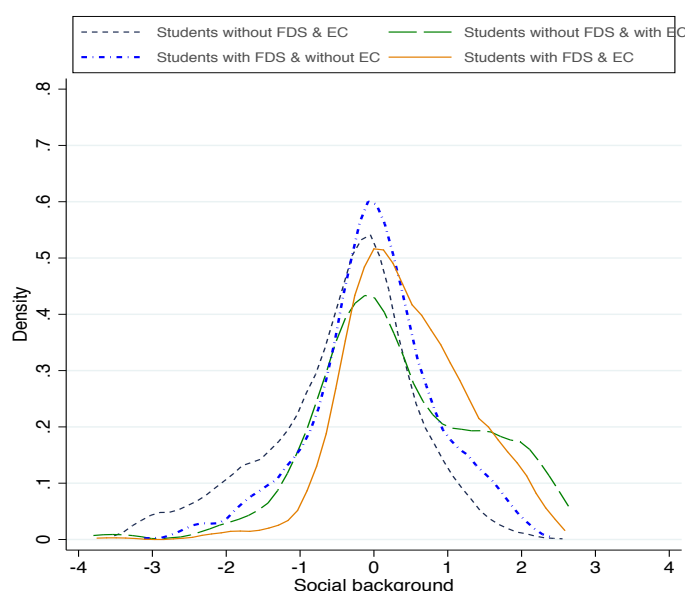
Source: Authors' calculations based on the Young Lives School Survey 2011

This is firstly because students from higher social backgrounds have better access to FDS (Table 2-2). Less than one-quarter of students from the most disadvantaged backgrounds (quintile 1) benefit from FDS. This proportion almost doubles for the 40 percent of students from the highest social background (quintiles 4 and 5). Thus, despite the implementation of the School Education Quality Assurance Program, students from low social backgrounds are still lagging behind in access to FDS. Poor SEQAP targeting may partially explain the persistence of low access to FDS for disadvantaged

children. Although this program targeted the most disadvantaged children, exclusion errors are substantial: our data show that 86 percent of students in the poorest quintile have not benefitted from SEQAP. Reversely, inclusion errors are significant: over half (57 percent) of SEQAP students are found in the three most advantaged social background quintiles (Table 2- 3).⁴⁷

The second reason for the gap in instructional time across social backgrounds is based on well-off students' attendance of extra classes. Children with a positive social background index, corresponding more or less to the two highest social background quintiles, are more likely to attend both extra classes and FDS than FDS only (Figure 2- 1). For these children, there is no substitution between FDS and extra classes as expected. At the bottom of the social ladder, the likelihood of attending half-day schooling only is the highest.⁴⁸ For the intermediary levels of social background (the third quintile), FDS without extra classes is the most common set-up.

Figure 2- 1: Distribution of student social backgrounds by schooling type



Source: Authors' calculations based on the Young Lives School Survey 2011

Thus, types of schooling differ widely across social backgrounds. As each type of schooling is associated with different instructional time, the gap in instructional time remains. Table 2- 4 shows that FDS helps reduce the students' lack of instructional time. Attending FDS secures students a 35 percent increase in instructional periods compared to half-day schooling (with neither FDS nor extra classes). However, extra classes give students longer instructional time than FDS. Extra classes increase instructional time by 48 percent, which is 13 percentage points more than for FDS.

⁴⁷ The School Survey asked principals whether their school or commune is or has ever been part of SEQAP.

⁴⁸ This does not rule out the fact that some children from the lowest quintile of social background benefit from extra classes. However, the attendance of extra classes is marginal for these children compared to children in the highest quintile (21 per cent compared to 63 per cent, see Table 2- 2).

Consequently, FDS and extra classes together provide the longest instructional time of 45 periods per week, which is 79 percent higher than just half-day schooling. So FDS only partially narrows the gap in instructional time between students who can afford to pay for extra classes and those who cannot.

Table 2- 4: Mean instructional time across schooling types, the entire sample and student social backgrounds (number of periods)

	Without and EC	FDS	Without and with EC	FDS	With FDS and without EC	With FDS and EC
Q1	25.1		35.7		33.9	42.1
Q2	25.4		35.4		34.5	45.1
Q3	25.5		34.6		34.5	45.8
Q4	25.6		37.5		34.3	44
Q5	25.7		41.4		34.7	46.6
All	25.4		37.6		34.4	45.4
<i>Obs</i>	<i>1213</i>		<i>728</i>		<i>788</i>	<i>535</i>

Source: Authors' calculations based on the Young Lives School Survey 2011

Table 2- 5: Number of schooling periods by subject by FDS status and social background

Quintiles of social background		Q1	Q2	Q3	Q4	Q5	All
# of periods in Vietnamese	No FDS	7.9	7.8	7.7	7.8	7.5	7.8
	FDS	8.6	9	9.1	9.4	9.9	9.3
	Diff.	0.7	1.2	1.4	1.6	2.4	1.5
# of periods in Mathematics	No FDS	5	5	5	4.9	4.7	4.9
	FDS	6.3	6.3	6.4	6.6	7.2	6.6
	Diff.	1.2	1.3	1.4	1.7	2.6	1.7
# of periods in all subjects	No FDS	25	25	25.3	25.5	26.2	25.4
	FDS	33.8	34.3	34.5	34.5	34.7	34.4
	Diff.	8.7	9.3	9.2	8.9	8.5	9

Source: Authors' calculations based on the Young Lives School Survey 2011

A third reason why instructional time still differs a great deal across social backgrounds could be due to the large variations in the amount of additional time included in FDS. Table 2- 4 examines the number of periods embedded in each type of schooling for each quintile of the social background index in order to assess whether time related to FDS is the same for children from different social backgrounds. It shows that this third explanation does not hold up. The number of periods for those attending FDS only is almost the same for each social background index quintile. Only the number of

periods related to extra classes increases with the level of social background. However, closer inspection finds that FDS secures fewer additional periods in Vietnamese and Mathematics for students from low social backgrounds than for children from high social backgrounds (

Table 2- 5). Quintile 1 students gain less than one period in Vietnamese from FDS, while quintile 5 students gain two and a half periods.

In sum, FDS narrows the differences in instructional time due to extra classes, but does not close the gap. Since well-off students have more access to both FDS and extra classes, students from low social backgrounds lag behind in the number of instructional hours.

5. Do students from different social backgrounds have the same access to full-day schooling resources?

Table 2- 6: Frequency of students with school resources – FDS vs. non-FDS (%)

<i>School facilities (%)</i>	No FDS	FDS		<i>Class facilities (%)</i>	No FDS	FDS	
Major repairs needed	29.9	12.1	***	Blackboard or whiteboard	99	100	***
Separate room for G5	96.8	97		Wall map	70.7	67.8	*
Have library	64.4	89.3	***	Teacher's staffroom	83.1	92.4	***
Computers for students	32.1	55.9	***	Teacher's desks	98.4	100	***
Internet for students	28.3	51.4	***	Sufficient electric lights	96	98.5	***
Electricity	95.1	100	***	Electric fan	87	97	***
Electricity working today	91.9	95.5	***	TV	4.1	15.1	***
Latrines for students	87.1	98.5	***	Video player or DVD	1	0	***
Separate latrines for boys/girls	83.6	87.9	***	Radio	3	0	***
Clean drinking water for students	5.8	32	***	Overhead projector	7	6	
				Computer	6.1	16.6	***
				Books other than textbooks	69.3	84.2	***
<i>Principal</i>	No FDS	FDS		<i>Teacher</i>	No FDS	FDS	
Years of education	19.8	20.5	***	Years of education	18.9	19.1	***
Years of working at school	25.9	25.5		Years of teaching	18.5	16.1	***
In-service training >15 days (%)	31.2	55.9	***	In-service training>5 days (%)	49	55.1	***

Note: ***, **, * Difference significant at the 1%, 5%, 10% level respectively

Source: Authors' calculations based on the Young Lives School Survey 2011

The introduction of FDS calls for more school resources to ensure that it is operational and can be translated into learning achievement. So, it is important to know whether FDS is indeed underpinned by more resources at school level and whether these resources are equally distributed by

socioeconomic background of the children. Additional resources concern mainly school infrastructures such as new or rehabilitated classrooms, class facilities and pay increments for teachers to cover the additional workload. SEQAP also provides teachers and school principals with training to improve teaching and management practices. Therefore, the definition of school resources is limited here to school infrastructure, class and school facilities, and teacher and principal skills and experience. Additional pay is already reflected in instructional time.

Table 2- 6 shows the higher level of school resources in almost all areas when FDS is available. Students with FDS have better access to facilities at both school and class level compared to students without FDS. This is especially true for infrastructure resources: the need for major school repairs is halved among students with FDS; the proportion of students with access to clean drinking water is more than five times higher with FDS, albeit still low, showing that clean drinking water is not a common standard among schools in Vietnam; and most FDS students have latrines. School and class facilities are also better for students attending FDS. The share of students with a library at school is 1.4 times higher for FDS students. Access to computers and the Internet is almost two times higher with FDS. Classes with FDS are more often equipped with modern facilities such as TVs and electric fans. Yet these facilities are not part of the FDS specifications and may reflect a self-selection process in the introduction of FDS rather than improved facilities due to FDS: schools with more resources and financial contributions from the families are more likely to have introduced FDS.

Skills measured by the principals and head teachers' length of education⁴⁹ and in-service training are higher for students with FDS than for others, but the difference is very slight. In addition, teachers and principals are no more experienced where FDS is implemented.

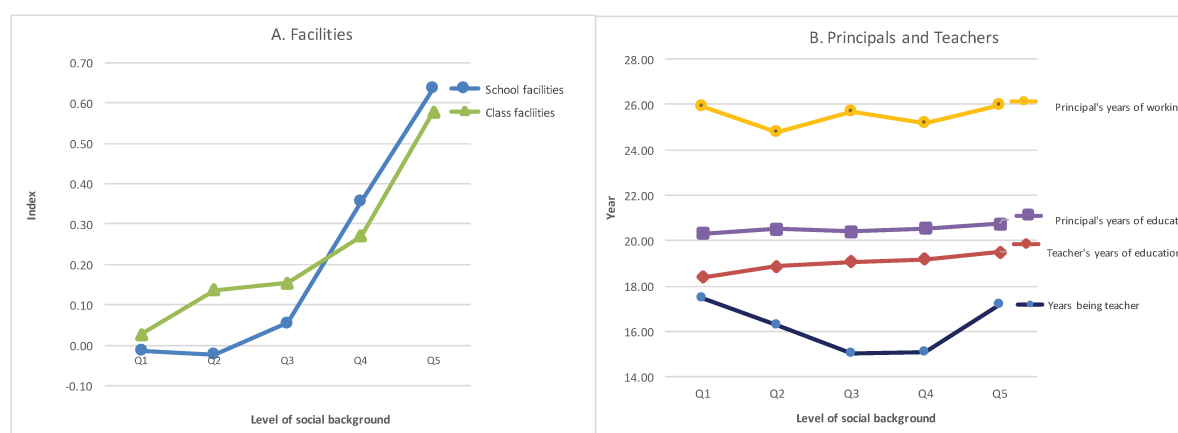
Thus, FDS students have access to better school resources than the non-FDS group. Yet are these resources equally distributed across FDS students from different social backgrounds? To answer this question, we create two indices synthesizing the information on school facilities and class facilities based on MCA analysis.⁵⁰ The higher the values of these indices, the more facilities to which the student has access in his/her school and class.

Figure 2- 2 shows the upward trend of school and class facilities across social backgrounds for FDS students: the higher a student's social background level, the higher the level of school and class facilities to which they have access. In addition, more advantaged social backgrounds also give students access to teachers with a higher level of education: on average, students in the lowest social background quintile have teachers with one year of education less than students in the highest social background quintile. On the other hand, no correlation is found either between teacher or principal experience and social background or between social background and the principal's level of education.

⁴⁹ The number of years of education refers to the sum of the number of years to obtain the highest level of general education completed and the number of years to obtain the highest level of teacher training qualification received.

⁵⁰ See detailed MCA results in Appendix A.4 to Appendix A.9.

Figure 2- 2: School resources by social background level for students attending FDS



Source: Authors' calculations based on the YLs School Survey 2011

These results show that students are sorted by social background into classes and schools with different levels of resources and, more importantly, that FDS does not counteract this selection process. When separately considering FDS implemented with SEQAP support, we find no correlation between the social background index and the school and class resources or the teacher's education.⁵¹ Hence, FDS with SEQAP support does not contribute to increase the resource gap and even may reduce it as school resources do not depend anymore on the social background of the students. However, only 18 per cent of students with FDS is in schools supported by SEQAP. This proportion is too small to offset the general trend towards selection process. Thus, FDS does not overcome inequality of opportunity in access to resources, inequality already observed by VASS (2012) and Hoang *et al.* (2012).

6. Does full-day schooling have a positive effect on learning achievement varying by social background?

This section tests whether students from disadvantaged backgrounds display sharper learning progress from attending FDS than the advantaged group. We estimate value-added models where students' test scores measured at the end of the academic year are regressed by the dummy of FDS attendance and a set of controls including the test scores at the beginning of the academic year. A finding of a reduction in inequality of opportunity in learning achievement implies a positive effect of FDS on learning progress that is higher for children from disadvantaged backgrounds, or if FDS has a detrimental effect on learning achievement, a less negative effect for children with lower social

⁵¹ Coefficients of correlation between social background index and facilities index or teacher's education rank between 0.02 and 0.08 and are not significant for students with FDS in SEQAP schools. These coefficients are significant and rank between 0.22 and 0.30 for students with FDS but in schools without SEQAP support.

backgrounds. Thus, we first estimate the effect of FDS on the entire population, and then the effect of FDS by social background level.

6.1. Estimation Strategy

We estimate a value-added model (VA) for the first step to test whether FDS has a positive and significant effect on learning progress. Value-added models were initially developed to identify the contribution of school and teacher characteristics to student achievement. These models are derived from a structural “education production function” in which learning achievement is modelled as a function of child, family and schooling inputs, and learning is viewed as both a dynamic and cumulative process. Empirical estimation of this function poses certain problems due to the lack of information on past inputs and the non-random assignment of students to schools and teachers. Value-added models circumvent these difficulties by assessing the difference in the learning achievement of students at the beginning and end of an academic year. Given that achievement is based on a combination of the student’s innate ability and all school inputs made available to the student through to the reference date, the model links current student learning outcomes to current educational inputs and lagged outcomes, which are assumed to capture all past educational inputs as well as the student’s innate ability. As a result, VA models have been largely used to deal with bias due to omitted past educational inputs and the children’s time-invariant unobserved characteristics such as their innate ability (Mizala & Romaguera, 2002; Todd & Wolpin, 2007; Guarino *et al.*, 2015).

Applying the review by Mizala & Romaguera (2002) and Todd & Wolpin (2007) of the VA model, a student i test score T_{it} in Math or Vietnamese is a function of a vector of current educational inputs X_{it} , a full-day schooling dummy E_{it} , lagged test score T_{it-1} and a residual e_{it} summing up the unobservable factors as follows:

$$T_{it} = \alpha X_{it} + \beta E_{it} + \gamma T_{i,t-1} + e_{it} \quad (1)$$

Cunha and Heckman (2009) additionally provide empirical evidence on the cross-productivity effects of skills. In other words, different skills substitute for each other. This feature implies ability in Math might contribute to ability in Vietnamese and vice versa. Consequently, equation (1) can be specifically expressed as:

$$T_{Mit} = \alpha_M X_{it} + \beta_M E_{it} + \gamma_M T_{Mi,t-1} + \theta_M T_{Vi,t-1} + e_M \quad (2)$$

$$T_{Vit} = \alpha_V X_{it} + \beta_V E_{it} + \gamma_V T_{Vi,t-1} + \theta_V T_{Mi,t-1} + e_V \quad (3)$$

in which subscripts M and V stand for Mathematics and Vietnamese respectively, and other symbols keep the same meaning as in equation (1).

Because the implementation of FDS is not random and may depend on unobserved school and student characteristics, the coefficients β_M and β_V of FDS might be biased. Potential bias comes first from unobserved school and student characteristics, factors that are correlated with both FDS and learning achievement. For example, highly motivated and highly able students are more likely to choose FDS schools and display better achievement. The same holds true at the school level if the schools that choose to implement FDS are, for example, the most dynamic schools or schools where parents are highly involved.

The introduction of the lagged test scores (test score at the beginning of the academic year) on the right-hand side of the equation controls for grouping and assignment mechanisms insofar as those grouping mechanisms are correlated with prior achievement (Guarino *et al.*, 2015). In other words, β_M and β_V coefficients are unbiased if two students with the same test scores obtain the same year-end test scores, regardless of their unobserved characteristics and school-related unobserved characteristics. Although this is a strong assumption, we expect potential bias should this assumption not hold to be mitigated by the introduction of school resources variables (school facilities, class facilities, teacher and principal skills, and teacher perception of the lack of resources) and social background variables. These variables are able to partially capture the school and student heterogeneity that could impact the learning dynamics between the test and year-end test.

Another source of bias may arise from self-selection due to the sampling design of the School Survey. More than one-third of the children surveyed (34.5 percent) are children included in the Young Lives longitudinal survey sample selected in 2002 when the children were one year old (hereafter called YL children). In addition, among the 2,000 YL children, only 56 percent of them were in Grade 5 in 2011 and therefore included in the School Survey sample. YL children are thus overrepresented in the School Survey and their selection in Grade 5 in 2011 is obviously not random. This is mostly due to late enrolment and repetition and hence correlates with the children's unobserved characteristics. The effect of FDS could be overestimated if this selection process is not taken into account. For this reason, we restrict the estimation of learning achievement to the sample of the 1,125 YL children surveyed in both the Young Lives longitudinal surveys and the School Survey in order to control for the selection process for these children. A Heckman two-step selection model is estimated (2SLS). The selection equation consists in estimating the probability for YL children surveyed in 2002 to be included in the School Survey or, in other words, to be in Grade 5 in 2011. We then estimate equations (2) and (3) by adding the inverse Mills ratio from the selection equation.

Despite our strategies to limit the self-selection bias due to the unobserved characteristics of the children and educational inputs, the idea that some bias remains cannot be completely rejected, especially biases related to time-varying unobserved characteristics, in common with other studies that do not use randomized trials in the implementation of instructional time-lengthening policies. Therefore, the coefficients β_M and β_V of FDS should be interpreted as the association between learning progress and FDS, and not as causal inference.

In keeping with Angrist and Krueger (1991), age in months in September 2011 is used in the selection equation. Age in months is clearly a predictor of selection as it determines late entry into school. Whether this variable satisfies the exclusion condition is debated in the literature. On the one hand, birth month is generally accepted as being random. On the other hand, some studies have demonstrated that the season of birth affects school achievement: younger children in the same age group display worse performance than the older children because of their lower level of maturity (see Verachtert *et al.*, 2010, for a review). However, the effect of birth month on achievement diminishes as children grow older. In addition, this effect should be captured in our estimation by the test scores at the beginning of the academic year and by the selection process, as birth month should affect the likelihood of repetition. For these reasons, it is reasonable to consider that age in months is uncorrelated with the residual term of equations (2) and (3).

Other independent variables included in the selection equation are child gender, a dummy that takes the value 1 if the child belongs to a minority ethnic group, the number of older siblings, location (urban/rural area), and two variables designed to capture social background. These variables are the highest grade completed by the caregiver at school and the Young Lives wealth index in Round 2 (when the child is aged 5), which is composed of three sub-indices in the form of a housing quality index, an access-to-services index, and an index of asset ownership.⁵²

In equations (2) and (3), vector X_{it} includes the student's characteristics, the characteristics of his/her family and current educational inputs. The student's characteristics controlled for are gender, ethnic group, number of siblings, number of older siblings, location (urban/rural area), and social background index. The importance of these variables to learning achievement has been identified across a range of contexts (Buchmann and Hannum, 2001). The variables used to control for current educational inputs are the class facilities index and the teacher's skills measured by the number of years of education and training.⁵³ We add the teacher's subjective perception about the lack of resources in his/her class in order to mitigate measurement error with respect to class resources. This is a dummy taking the value 1 if the teacher considers the lack of resources (e.g. textbooks and materials) to be a problem for student learning in his/her class. We also control for extra-class attendance as it may be a substitute for FDS in terms of instructional time. Extra-class attendance is introduced by means of a dummy variable rather than the number of hours in order to limit its collinearity with the social background index.⁵⁴ Descriptions of the variables are provided in Appendix A.10.

The 2SLS estimation results for equations (2) and (3) are shown in Table 2- 7 for the effect of FDS and in the appendix for the results as a whole (Appendix A.12 for Mathematics and 0A.13 for Vietnamese). Model (1) is the estimation for the whole sample. The estimation of the selection equation is presented in Appendix A.11.

The heterogeneity of the effect of FDS on student achievement across students from different social background levels is addressed by estimating equations (2) and (3) by subgroup of social backgrounds. Each group corresponds to one quartile of the social background index. The corresponding models are labelled from Model (2) for the lowest social background quartile to (5) for the highest quartile.

6.2. Estimation Results

⁵² See Nguyen (2008) for more details on the wealth index. Social background cannot be measured in the same way in the selection equation as in equation (2) and (3) as the set of variables differs between the Young Lives longitudinal surveys and the School Survey.

⁵³ The school facilities index is not included to avoid collinearity issues with FDS variables.

⁵⁴ Only extra classes in Vietnamese and Mathematics are considered here.

For both Math and Vietnamese, attending FDS does not equate with higher learning progress *ceteris paribus*. In Models (1), FDS attendance is even negatively and significantly associated with learning progress. Compared to students attending FDS, students not in FDS but with same observable characteristics display 16 and 24 points more in their year-end test scores in Mathematics and Vietnamese respectively (Table 2- 7, Models (1)).

This result goes against Tran (2014)'s findings who concludes to an improvement of achievements due to the implementation of FDS. However, her results are based on the comparison of the share of good students before and after the implementation of the FDS in six schools only, and the good students are defined according to non-standardized tests developed at the school level. Beside the non-representativeness of her sample, the teachers especially those benefiting from the SEQAP may be incited to over-evaluate the share of good students in order to demonstrate their capacity to use the support of the government efficiently.

Non-significant or negative effect of the lengthening of instructional time is found in other contexts. By way of illustration, Xerxenevsky (2012) finds no effect of the Brazilian program of extended day *Mais Educação* for both mathematics and language in eighth grade, and even a significant negative effect for mathematics in fourth grade. While most of the studies find that the lengthening of instructional time result in positive increases in achievement, the effect is small. One exception is Orkin (2013) who find large positive effect of FDS in the first three grades of schools. However, the measure of achievement at her disposal is not very precise.⁵⁵ By the way, studies are difficult to compare. First, they evaluate the effect on different grades, including primary and secondary education. Pedagogical approaches vary tremendously across grades, making the generalization of the results to other grades limited. Second, the design of the programs differs from one study to another in the size of the extension of time but also in the accompanying measures to the lengthening of the instructional time. Third, the availability of the data induces various methodological approaches to identify the effect, exploiting variations in instructional time across students, across subject (within-student variations) or using propensity score matching when longitudinal data are not available. Consequently, Holland *et al.* (2015) conclude their literature review on the extension of the school day in Latin America and the Caribbean by saying, *"Although we have insufficient data to explain precisely why one might find this range of positive, negative, and insignificant effects, they certainly demonstrate that while there may be a tendency toward positive academic outcomes, this is certainly not guaranteed."*

One reason for finding the negative effect of FDS could reside in the fact that learning progress is measured by means of Mathematics and Vietnamese tests. The increase in instructional time due to FDS is minimal these subjects, at just 1.5 and 1.7 additional periods per week in Vietnamese and Mathematics respectively, and these two subjects account for around one-third of the total increase in instructional time (

⁵⁵ Her measure of achievement in Mathematics is the likelihood to be numerate and in language to be able to write. She finds no significant effect on the likelihood to be able to read.

Table 2- 5). Additional hours are mostly used for non-academic subjects. This practice could lead to a detrimental effect on Mathematics and Language achievements if additional non-academic hours displace time that would be spent on homework in these two subjects. The data show some evidence in line with this assumption: around half of the children with FDS declare that they rarely or never have Vietnamese homework and this proportion is only one-third for students without FDS. For Mathematics, the proportions are respectively 40 and 27 per cent. Another side effect of FDS on learning progress could be that children may lose concentration by spending all day at school if the curricula are not adapted to this change or if they are less able to attend school when it requires them to spend all day at school instead of helping their parents in their daily tasks. However, the data do not enable us to investigate further the channels through which FDS negatively affects children progresses.

Table 2- 7: Estimation results of learning progress with 2SLS (Beta coefficients)

	(1)	(2)	(3)	(4)	(5)
	All	Q1	Q2	Q3	Q4
Models for retest scores in Mathematics					
FDS attendance	-16.37***	-14.38	-25.33	-5.688	-19.98
	(5.151)	(11.62)	(13.91)	(56.77)	(21.08)
Models for retest scores in Vietnamese					
FDS attendance	-23.90***	-57.45**	-19.27	-5.498	-25.36
	(5.631)	(26.96)	(4.966e+08)	(12.96)	(30.99)
Observations	1,053	1,053	269	269	273

Note: Standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Source: Authors' calculations based on Young Lives School Survey 2011 and Household Survey 2006.

However, one can object that the negative effect of FDS is due to heterogeneity in the implementation of FDS across schools, which we are unable to control for. If FDS is mostly implemented in low performing schools, it is possible that negative coefficients reflect inadequate controlling for pre-existing lower trajectories of student performance in FDS schools. It cannot be ruled out as the SEQAP targets the schools with the most disadvantaged children. But we can test it by distinguishing in equations (2) and (3) between FDS implemented in SEQAP schools and FDS in schools without governmental support. If the negative effect of FDS results from a non-random selection process in the implementation of FDS, we would expect that the effect of FDS is negative for SEQAP children who are on average with more disadvantaged social background and have lower achievement at the start of the year. Reversely, we expect a positive effect for FDS students in schools

without governmental support as their higher social background, their higher level of achievement and level of resources at their disposal may reflect unobserved upward learning trajectories.

Table 2- 8 presents the results of the estimation of the equations (2) and (3) where the dummy for FDS is replaced by a categorical variable which takes the value 0 if no FDS, the value 1 if the FDS is implemented with SEQAP support and the value 2 if the FDS is implemented without SEQAP support. It shows that the negative effect of FDS on Mathematics learning progress is mostly driven by the effect of FDS implemented out of SEQAP. Without ruling it completely out, this does not support the hypothesis that our results are driven by unobserved selection of schools in the FDS process. For Vietnamese learning achievement, it is less clear as both types of FDS have negative and significant effect. However, we would expect a neutral or positive effect of FDS without SEQAP if an unobserved selection process would be at play. We conclude therefore to a detrimental effect of FDS on learning progress.

Table 2- 8: Estimation results of learning progress with 2SLS by distinguishing between FDS with and without SEQAP (Beta coefficients)

	(6)	(7)	(8)	(9)	(10)
	All	Q1	Q2	Q3	Q4
Models for year-end test scores in Mathematics					
FDS with SEQAP	-12.07 (8.874)	-2.550 (20.81)	-38.76 (24.60)	-2.400 (93.15)	-4.533 (37.23)
FDS without SEQAP	-17.30*** (5.387)	-16.81 (12.17)	-21.66 (15.29)	-6.353 (58.71)	-22.95 (22.26)
Models for year-end test scores in Vietnamese					
FDS with SEQAP	-34.34*** (9.734)	-62.85 (49.98)	-43.26 (1.015e+09)	-23.18 (19.23)	-11.22 (54.34)
FDS without SEQAP	-21.66*** (5.880)	-56.41** (28.11)	-13.12 (6.117e+08)	-1.987 (11.90)	-28.51 (34.30)
Observations	1,053	269	269	273	265

Note: Standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Source: Authors' calculations based on Young Lives School Survey 2011 and Household Survey 2006.

In a second step, we examine whether this effect is more detrimental for well-off student and thus whether the FDS participate to the reduction of the learning achievement gap. In fact, the opposite is true for learning Vietnamese. As shown in Table 2- 7 and Table 2- 8, children in the 25 percent of

families from the lowest social backgrounds display slower progress in Vietnamese when they benefit from FDS than when they do not (Models (2) to (5)) and this is mostly driven by the FDS without SEQAP (Models 6). For these children, FDS may be of poorer quality as shown by the smallest increase in Vietnamese instruction periods for these children than for others (

Table 2- 5), and the substitution of homework time in Vietnamese by non-academic hours may have a more serious consequence for them than for other children. However, we cannot test this assumption properly with our data. However, we find no differential effect of FDS on Mathematics learning achievement across quartiles of social background

To sum up, these results show that FDS did not provide students from low social backgrounds with enough or suitable additional instructional time to offset their learning progress disadvantage compared to better-off students. Worse still, FDS may contribute to inequality of opportunity in education, because its implementation conditions appear to be more detrimental to students from low social backgrounds for learning progress in Vietnamese.

6.3. Robustness checks

In this section, we test whether these results are robust to the specification of FDS and to the choice of the sample. As mentioned in section 3, FDS could be defined based on the statement made by student or teacher. FDS attendance as reported by teacher takes the value 1 if teacher answers that all pupils attend FDS when asked, *“Do the pupils in your class attend FDS?”* With respect to the student’s statement, the FDS dummy takes the value 1 if the student answers yes to the question, *“Do you attend full-day schooling (two shifts- morning and afternoon)?”* Although these measures are not our favourite ones for the reasons developed in section 3, it is interesting to test whether our results hold true when considering these two alternative definitions. We also use the number of periods the pupils have in the week according to teacher’s statements as an alternative measure. This reflects the instructional time at school instead of the FDS implementation, but its effect should be consistent with those of FDS if the relationship between instructional time and achievements is linear.

Models (2) to (6) are estimated using these alternative measures. Results are presented in Table 2- 9 where models (11) are based on 2SLS estimation across all YL children included in the School Survey, and models (12) to (15) are for the subgroups defined by social background quartile.

Another way to test the robustness of our result is to estimate the learning progress model for the entire School Survey sample, i.e. without taking into account the sampling selection bias due to the over-representation of YL children in the sample. Value-added models are estimated with OLS and results are presented in Table 2- 10. Model (16) covers the entire School Survey sample while models (17) to (20) are on the subsample defined social background quartile.

Again, we find no positive effect of FDS or number of periods on learning progress in Mathematics and Vietnamese irrespective of the sample or measurement of FDS. FDS has a significant and negative effect on Mathematics learning progress using either 2SLS estimation or OLS estimation when the entire population is considered, except Model (11) where FDS teacher’s statement is used and no significant effect is found. This negative effect holds true when quartiles of the social background are

distinguished in the OLS estimation (Models (16) to (20)). As for the main models, this suggests that there is no differential effect of FDS across the social background in Mathematics learning achievement.

As far as learning progress in Vietnamese is concerned, a negative and significant effect of FDS is found again for the population as a whole. However, results on the differential effect across social background are less robust to the specifications. The more detrimental effect of FDS on learning progress in Vietnamese for students from the lowest social backgrounds is observed in just two of three specifications. Yet this negative and significant effect is detected for the other quartile of social background as well in two of three specifications, albeit to a lesser extent.

The number of periods has a negative and significant effect as well (Model (11)), and more for the students from the lowest social backgrounds (Model (12)) but for Vietnamese learning achievement only. Considering that the number of periods does not include Vietnamese teaching only, this shows that additional instructional time does not guarantee improvement in learning achievement and even may have a detrimental effect.

Thus, the robustness checks confirm that FDS schooling do not contribute to reducing educational gap. Whether it contributes or not to an increase in inequality is less clear.

Table 2- 9: Estimation results of learning progress with 2SLS with alternative measures of FDS (Beta coefficients)

	(11) All	(12) Q1	(13) Q2	(14) Q3	(15) Q4
Models for retest scores in Mathematics					
FDS- teacher's statement	-7.113 (4.894)	-7.509 (10.05)	-13.54 (13.80)	15.31 (56.06)	-23.00 (21.46)
FDS- student's statement	-13.78*** (5.194)	-18.19 (13.07)	-12.15 (13.59)	3.708 (64.27)	-31.02 (20.35)
Number of periods	-0.854* (0.494)	-1.573 (1.144)	-1.784 (1.190)	0.947 (5.240)	-2.309 (2.195)
Models for retest scores in Vietnamese					
FDS- teacher's statement	-14.29*** (5.349)	-7.373 (15.47)	-16.03 (10.52)	6.202 (14.61)	-20.20 (20.32)
FDS- student's statement	-24.25*** (5.676)	-32.48** (14.14)	-21.82** (11.04)	-1.711 (14.94)	-23.88* (13.74)
Number of periods	-1.633*** (0.540)	-6.052* (3.133)	-1.028 (1.026)	0.190 (1.252)	-2.228 (3.538)

Observations	1,053	269	269	273	265
--------------	-------	-----	-----	-----	-----

Note: Standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Source: Authors' calculations based on Young Lives School Survey 2011 and Household Survey 2006.

Table 2- 10: Estimation results of learning progress with OLS on the entire School Survey sample
(Beta coefficients)

	(11) All	(12) Q1	(13) Q2	(14) Q3	(15) Q4
Models for retest scores in Mathematics					
FDS- Teacher's statement	-7.113 (4.894)	-7.509 (10.05)	-13.54 (13.80)	15.31 (56.06)	-23.00 (21.46)
FDS- Student's statement	-13.78*** (5.194)	-18.19 (13.07)	-12.15 (13.59)	3.708 (64.27)	-31.02 (20.35)
Number of periods	-0.854* (0.494)	-1.573 (1.144)	-1.784 (1.190)	0.947 (5.240)	-2.309 (2.195)
Models for retest scores in Vietnamese					
FDS- Teacher's statement	-14.29*** (5.349)	-7.373 (15.47)	-16.03 (10.52)	6.202 (14.61)	-20.20 (20.32)
FDS- Student's statement	-24.25*** (5.676)	-32.48** (14.14)	-21.82** (11.04)	-1.711 (14.94)	-23.88* (13.74)
Number of periods	-1.633*** (0.540)	-6.052* (3.133)	-1.028 (1.026)	0.190 (1.252)	-2.228 (3.538)
Observations	1,053	269	269	273	265

Note: Standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Source: Authors' calculations based on Young Lives School Survey 2011 and Household Survey 2006.

7. Conclusion

FDS was introduced initially to address a lack of instructional time. The lengthening of instructional time brought about by FDS should drive a higher quality of education. FDS may also help to restrict the negative effects of widespread extra classes in Vietnam on the inequality of opportunity in education. However, these two objectives have not yet been achieved: learning progress is not better with FDS and a large gap remains between children from different social backgrounds even though they benefit from FDS.

Vietnam's particularity is that FDS has been introduced in a range of configurations to adapt to the schools' physical and human resources. These resources rely mainly on the communities and families' means. Consequently, there is still a huge gap in instructional time across social backgrounds:

students from low social backgrounds have less access to FDS and, when they do benefit, they still have much less instructional time than more advantaged students. This is because FDS gives students less instructional time than extra classes and because students from advantaged social backgrounds generally attend both FDS and extra classes. In addition, the increase in instructional time in Mathematics and Vietnamese due to FDS is not equally distributed across students, and students from low social backgrounds are disadvantaged in this respect. Secondly, student access to well-equipped school and class facilities and skilled, experienced teachers rises with their level of social background. Briefly speaking, FDS students from low social backgrounds have less access to school resources compared with their counterparts from more advantaged social backgrounds who benefit from FDS.

As a result, there is no evidence of a positive relationship between FDS and learning progress. FDS is not associated with better learning progress, especially for children from disadvantaged backgrounds. Even FDS seems to have a detrimental effect of learning achievements. At this stage of implementation, FDS does not appear to have reduced the inequality of opportunity in learning achievement. In fact, it seems to have had the opposite effect by magnifying the effect of social background on learning progress, thereby widening the gap in learning achievement in Vietnamese.

The above evidence provides policy pointers for FDS implementation to achieve the target of educational equality in Vietnam. First of all, FDS support should be better targeted for students from disadvantaged backgrounds to offset their gaps in instructional time compared to students from higher social backgrounds. Secondly, it is important to introduce more coherent FDS regulations. The regulations should guarantee the same instructional time to all students, with a subsequent number of additional academic time. Furthermore, additional school resources, training and tailored curricula should be provided to limit the detrimental side effects of increasing instructional time, especially for disadvantaged children.

This paper contributes to the literature on the impact of longer instructional time by demonstrating a negative effect of longer instructional time in a context where the extension of instructional time receives poor government support, families make a substantial contribution to funding the policy, and where the standard of this extension of time are poorly defined.

Acknowledgements

This paper has received funding from the Privatisation in Education Research Initiative (PERI, <http://www.periglobal.org/>) and from the NOPOOR project (www.nopoor.eu) under the European Commission's FP7. The authors would like to thank anonymous reviewers for their comments on earlier drafts of the paper.

References

- Angrist, D., Krueger A.B. (1991). Does Compulsory School Attendance Affect Schooling and Earnings? *The Quarterly Journal of Economics*, vol.106(4), pp.979-1014.
- Arzola González, M. P. (2010). Impacto de la Jornada Escolar Completa en el Desempeño de los Alumnos, medido con la Evolución en sus Pruebas SIMCE. Santiago: Pontificia Universidad Católica de Chile.
- Biernat, L. (2011). Reducing the Achievement Gap: There is Only One Real Solution. *SSRN Electronic Journal*. <http://doi.org/10.2139/ssrn.1936041>
- Buchmann, C., Hannum, E. (2001). "Education and Stratification in Developing Countries: A Review of Theories and Research", *Annual Review of Sociology*, vol.27, pp.77-102.
- Bodewig, C., Badiani-Magnusson, R., Macdonald, K., Newhouse, D., Rutkowski J. (2014). *Skilling up Vietnam: preparing the workforce for a modern market economy*. Washington, DC: World Bank.
- Cabrera-Hernandez, F. (2015). Does lengthening the school day increase students' academic achievement? Evidence from a natural experiment. Working Paper Series No. 74-2015, Department of Economics, University of Sussex. Retrieved from <http://www.sussex.ac.uk/economics/documents/wps-74-2015.pdf>
- Cunha, F., & Heckman, J. (2009). *The Economics and Psychology of Inequality and Human Development* (No. w14695). Cambridge, MA: National Bureau of Economic Research. Retrieved from <http://www.nber.org/papers/w14695.pdf>
- Dang, H.-A. (2007). The determinants and impact of private tutoring classes in Vietnam. *Economics of Education Review*, 26(6), 683–698. <http://doi.org/10.1016/j.econedurev.2007.10.003>
- Embretson, S. E., & Reise, S. P. (2000). *Item response theory for psychologists*. Mahwah, N.J: L. Erlbaum Associates.
- Guarino, C. M., Reckase, M. D., Stacy, B., Wooldridge, J. M. (2015). A comparison of Student Growth Percentile and Value-Added Models of Teacher Performance, *Statistics and Public Policy*, 2:1, 1-11, DOI: 10.1080/2330443X.2015.1034820
- Hincapié D. (2016). Do Longer School Days Improve Student Achievement? Evidence from Colombia. *Inter-American Development Bank Working Paper* 679.
- Hoang, X. T., Nguyen, T. P., Vu, V. N., Do, T. Q., Nguyen, T. H., Dang, T. H., & Nguyen, T. G. (2012). *Inequality Perception Study in Vietnam*". Background paper for the 2012 Poverty Assessment.
- Holland P., Alfaro P., Evans K. (2015) *Extending the School Day in Latin America and the Caribbean*. World Bank Policy Research Working Paper 7309, <http://dx.doi.org/10.1596/1813-9450-7309>

- Lavy, V. (2010). Do differences in school instruction time explain international achievement gaps in math, science and reading? Evidence from developed and developing countries. National Bureau of Economic Research Working Paper 16227. Retrieved from <http://www.nber.org/papers/w16227.pdf>
- Lavy, V. (2012). Expanding school resources and increasing time on task: Effects of a policy experiment in Israel on student academic achievement and behavior. National Bureau of Economic Research Working Paper 18369. Retrieved from <http://www.nber.org/papers/w18369>
- Mizala, A., & Romaguera, P. (2002). Equity and Educational Performance. *Economía*, 2(2), 219–273. <http://doi.org/10.1353/eco.2002.0009>
- Nguyen P. Ngoc (2008) An assessment of the Young Lives Sampling Approach in Vietnam, Young Lives Technical Note n°4.
- Orkin, K. (2013). The Effect of Lengthening the School Day on Children's Achievement in Ethiopia. Young Lives Working Paper 119. Retrieved from http://www.younglives-ethiopia.org/files/working-papers/wp119_orkin_ethiopia
- Patall, E. A., Cooper, H., & Allen, A. B. (2010). Extending the School Day or School Year: A Systematic Review of Research (1985-2009). *Review of Educational Research*, 80(3), 401–436. <http://doi.org/10.3102/0034654310377086>
- Schutz, G., Ursprung H., & Wossman L. (2008). Education Policy and Equality of Opportunity. *KYLOS*, 61(8): 279-308.
- Roemer, J. E. (1998). Equality of opportunity. Cambridge, Mass.: Harvard Univ. Press.
- Rolleston, C., James, Z., Tran, N. T. M. T., & Pasquier-Doumer, L. (2013). Making Progress: Report of the Young Lives School Survey in Vietnam. Young Lives Working Paper 100. Retrieved from http://www.younglives.org.uk/publications/WP/making-progress-report-of-the-young-lives-school-survey-in-vietnam/yl-wp100_rolleston
- Todd, P. E., & Wolpin, K. I. (2007). The Production of Cognitive Achievement in Children: Home, School, and Racial Test Score Gaps. *Journal of Human Capital*, 1(1), 91–136. <http://doi.org/10.1086/526401>
- Tran, N. T. M. T., & Pasquier-Doumer, L. (2015). Does full-day schooling increase educational inequality in Vietnam? PERI Young Lives ESP Working Paper Series 2015 No. 72.
- Tran, T. B. L. (2014). Full-day schooling Performance of Primary Schools in Disadvantaged Areas in Vietnam: A Comparative Case Study. *VNU Journal of Sciences: Education Research*, 30(No. 4), 17–30.
- VASS. (2012). Opportunities for Children in Vietnam. (Vietnam Academy of Social Sciences) Background paper for the 2012 Poverty Assessment.
- Verachtert, P., De Fraine, B., Onghena P., Ghesquiere P. (2010) Season of birth and school success in the early years of primary education, *Oxford Review of Education*, vol.36(3), pp.285-306.
- World Bank. (2011). Viet Nam: High Quality Education for All by 2020 (Vol. 2): Analytical report (Working paper No. 68092).
- Xerxenevsky L. L. (2012). Programa Mais Educação: Avaliação do Impacto da Educação Integral no Desempenho de Alunos no Rio Grande do Sul. Porto Alegre: PUCRS. Retrieved from <http://repositorio.pucrs.br:8080/dspace/handle/10923/2566>

Appendix

A. Tables

A.1. Descriptions of variables used for the social background index

Variable name	Category	Proportion (%)
Mother's educational level	Primary or less	53.12
	Lower secondary	23.15
	Upper secondary	23.73
Father's educational level	Primary or less	52.02
	Lower secondary	20.73
	Upper secondary	27.25
Number of meals per day	1	0.31
	2	16.57
	3 or more	83.12
<i>Home belongings</i>		
A television at home	No	5.39
	Yes	94.61
A radio at home	No	73.45
	Yes	26.55
An electric fan at home	No	10.93
	Yes	89.07
A mobile phone at home	No	8.65
	Yes	91.35
A bicycle at home	No	14.40
	Yes	85.60
A motorcycle at home	No	10.99
	Yes	89.01
A car at home	No	91.93
	Yes	8.07
Air conditioner at home	No	88.92
	Yes	11.08
A study lamp at home	No	27.83
	Yes	72.17
A study desk at home	No	12.33
	Yes	87.67
A study chair at home	No	12.70
	Yes	87.30
A pocket calculator at home	No	68.33
	Yes	31.67
Own place to study at home	No	20.20
	Yes	79.80
Number of books in the home	None	20.68
	1 to 5	20.28
	6 to 10	13.57
	More than 10	45.47
A computer at home	No	77.13
	Yes	22.87
Internet at home	No	81.79
	Yes	18.21

Source: Authors' calculations based on Young Lives School Survey 2011.

A.2. MCA for Social background index

Dimension	Principal inertia	Percent	Cumulation percent
dim1	0.033287	78.31	78.31
dim2	0.003861	9.08	87.40
dim3	0.000396	0.93	88.33
dim4	0.000127	0.30	88.62
dim5	0.000058	0.14	88.76
dim6	0.000005	0.01	88.77
dim7	0.000001	0.00	88.78
dim8	0.000000	0.00	88.78
Total	0.042505	100.00	

Source: Authors' calculations based on Young Lives School Survey 2011.

A.3. MCA Statistics of columns – Social background index

		Overall			Dimension 1		
		mass	quality	%inert	coord	sqcorr	contrib
Mother's educational level	Primary or less	0.0279	0.7307	0.0216	0.8489	0.7302	0.0201
	Lower secondary	0.0122	0.4128	0.0093	-0.0537	0.0030	0.0000
	Upper secondary	0.0125	0.8274	0.0441	-1.8465	0.7553	0.0425
Father's educational level	Primary or less	0.0274	0.7056	0.0203	0.8179	0.7053	0.0183
	Lower secondary	0.0109	0.3878	0.0086	0.1328	0.0176	0.0002
	Upper secondary	0.0144	0.8182	0.0404	-1.6595	0.7660	0.0395
Number of meals per day	1	0.0002	0.3561	0.0004	1.1020	0.3559	0.0002
	2	0.0087	0.9319	0.0059	0.8832	0.8948	0.0068
	3 or more	0.0438	0.9314	0.0012	-0.1792	0.8961	0.0014
A television at home	No	0.0027	0.9232	0.0338	3.5435	0.7996	0.0345
	Yes	0.0499	0.9232	0.0019	-0.1951	0.7996	0.0019
A radio at home	No	0.0387	0.9904	0.0061	0.4399	0.9659	0.0075
	Yes	0.0140	0.9904	0.0168	-1.2181	0.9659	0.0207
An electric fan at home	No	0.0057	0.9672	0.0294	2.3915	0.8672	0.0325
	Yes	0.0469	0.9672	0.0036	-0.2899	0.8672	0.0039
A telephone at home	No	0.0045	0.9145	0.0214	2.2753	0.8509	0.0233
	Yes	0.0481	0.9145	0.0020	-0.2125	0.8509	0.0022
A bicycle at home	No	0.0076	0.9547	0.0195	1.5931	0.7710	0.0192
	Yes	0.0451	0.9547	0.0033	-0.2676	0.7710	0.0032
A motorcycle at home	No	0.0057	0.9067	0.0238	2.1655	0.8823	0.0268
	Yes	0.0469	0.9067	0.0029	-0.2642	0.8823	0.0033
A car at home	No	0.0484	0.9067	0.0021	0.1983	0.7088	0.0019
	Yes	0.0042	0.9067	0.0240	-2.2658	0.7088	0.0217
Air conditioner at home	No	0.0467	0.8952	0.0073	0.3865	0.7530	0.0070
	Yes	0.0059	0.8952	0.0573	-3.0500	0.7530	0.0551
A study lamp at home	No	0.0146	0.9363	0.0505	1.9943	0.9008	0.0581
	Yes	0.0380	0.9363	0.0194	-0.7668	0.9008	0.0224
A study desk at home	No	0.0064	0.8423	0.0759	3.3351	0.7338	0.0711
	Yes	0.0462	0.8423	0.0105	-0.4611	0.7338	0.0098
A study chair at home	No	0.0066	0.8621	0.0670	3.1183	0.7503	0.0642
	Yes	0.0460	0.8621	0.0096	-0.4475	0.7503	0.0092
A pocket calculator	No	0.0358	0.9758	0.0142	0.6911	0.9462	0.0171
	Yes	0.0168	0.9758	0.0302	-1.4730	0.9462	0.0365
Own place to study at home	No	0.0107	0.9234	0.0477	2.2064	0.8542	0.0520
	Yes	0.0420	0.9234	0.0121	-0.5615	0.8542	0.0132
Number of books at home	None	0.0109	0.9778	0.0305	1.8304	0.9357	0.0364
	1 to 5	0.0106	0.8872	0.0032	0.3556	0.3238	0.0013
	6 to 10	0.0072	0.6487	0.0014	-0.1655	0.1099	0.0002
	More than 10	0.0240	0.9779	0.0173	-0.9381	0.9563	0.0211
A computer at home	No	0.0405	0.8434	0.0241	0.7595	0.7601	0.0233
	Yes	0.0121	0.8434	0.0802	-2.5308	0.7601	0.0778
Internet at home	No	0.0430	0.8339	0.0182	0.6273	0.7277	0.0169
	Yes	0.0096	0.8339	0.0811	-2.7951	0.7277	0.0754

Source: Authors' calculations based on Young Lives School Survey 2011.

A.4. MCA for School facilities index

Dimension	Principal inertia	Percent	Cumulation percent
dim1	0.043530	62.92	62.92
dim2	0.007214	10.43	73.35
dim3	0.000758	1.10	74.44
dim4	0.000030	0.04	74.49
Total	0.069183	100.00	

Source: Authors' calculations based on Young Lives School Survey 2011.

A.5. MCA Statistics of columns - School facilities index

		Overall			Dimension 1		
		mass	quality	%inert	coord	sqcorr	contrib
Major repairs needed	No	0.0814	0.8740	0.0036	0.1804	0.4688	0.0026
	Yes	0.0186	0.8740	0.0156	-0.7906	0.4688	0.0116
Separate room for G5	No	0.0022	0.6273	0.0207	-2.8603	0.5569	0.0183
	Yes	0.0978	0.6273	0.0005	0.0656	0.5569	0.0004
Have library	No	0.0230	0.8092	0.1206	-2.5556	0.7846	0.1504
	Yes	0.0770	0.8092	0.0361	0.7646	0.7846	0.0450
Computers for students	No	0.0580	0.7155	0.0955	-1.3108	0.6567	0.0997
	Yes	0.0420	0.7155	0.1320	1.8113	0.6567	0.1377
Internet for students	No	0.0601	0.7197	0.0878	-1.2377	0.6592	0.0920
	Yes	0.0399	0.7197	0.1320	1.8608	0.6592	0.1383
Electricity	No	0.0008	0.5742	0.0434	-5.6986	0.3842	0.0265
	Yes	0.0992	0.5742	0.0004	0.0468	0.3842	0.0002
Working electricity today	No	0.0042	0.6756	0.0535	-3.1637	0.4958	0.0422
	Yes	0.0958	0.6756	0.0024	0.1391	0.4958	0.0019
Latrines for students	No	0.0028	0.7740	0.0611	-4.7584	0.6570	0.0638
	Yes	0.0972	0.7740	0.0018	0.1380	0.6570	0.0019
Separate latrines for boys/girls	No	0.0117	0.7851	0.1270	-3.5429	0.7290	0.1471
	Yes	0.0883	0.7851	0.0169	0.4703	0.7290	0.0195
Clean drinking water for students	No	0.0817	0.6327	0.0091	0.0430	0.0105	0.0002
	Yes	0.0183	0.6327	0.0403	-0.1915	0.0105	0.0007

Source: Authors' calculations based on Young Lives School Survey 2011.

A.6. School facilities index by facility

Facility	No	Yes
Major repairs needed	-0.0087	-0.6770
Separate room for G5	-2.3537	-0.0924
Have library	-1.6952	0.3552
Computers for students	-0.9065	0.8829
Internet for students	-0.8651	0.9841
Electricity	-3.8004	-0.0540
Working electricity today	-2.5800	0.0094
Latrines for students	-2.7507	0.0675
Separate latrines for boys/girls	-1.5159	0.2523
Clean drinking water for students	-0.1639	-0.1592

Source: Authors' calculations based on Young Lives School Survey 2011.

A.7. MCA for Class facilities index

Dimension	Principal inertia	Percent	Cumulation percent
dim1	0.013994	50.60	50.60
dim2	0.003327	12.03	62.63
dim3	0.001765	6.38	69.01
dim4	0.000013	0.05	69.06
dim5	0.000000	0.00	69.06
Total	0.040056	100.00	

Source: Authors' calculations based on Young Lives School Survey 2011.

A.8. MCA Statistics of columns - Class facilities index

		Overall			Dimension 1		
		mass	quality	%inert	coord	sqcorr	contrib
Black board or white board	No	0.0005	0.1064	0.0304	1.1121	0.0105	0.0006
	Yes	0.0828	0.1064	0.0002	-0.0068	0.0105	0.0000
Wall map	No	0.0252	0.8983	0.0345	-1.5174	0.8529	0.0581
	Yes	0.0581	0.8983	0.0150	0.6596	0.8529	0.0253
Teacher's cabinet	No	0.0109	0.8406	0.0976	-3.8067	0.8157	0.1574
	Yes	0.0725	0.8406	0.0146	0.5705	0.8157	0.0236
Teacher's desks	No	0.0008	0.7895	0.0669	-10.5465	0.6831	0.0903
	Yes	0.0825	0.7895	0.0007	0.1038	0.6831	0.0009
Sufficient electric lights	No	0.0025	0.7803	0.1174	-8.1539	0.7057	0.1636
	Yes	0.0809	0.7803	0.0036	0.2482	0.7057	0.0050
Electric fan	No	0.0074	0.8269	0.0696	-3.8683	0.8086	0.1113
	Yes	0.0759	0.8269	0.0068	0.3789	0.8086	0.0109
TV	No	0.0762	0.4830	0.0121	-0.3096	0.3067	0.0073
	Yes	0.0071	0.4830	0.1293	3.3218	0.3067	0.0784
Video player or DVD	No	0.0828	0.1064	0.0002	-0.0068	0.0105	0.0000
	Yes	0.0005	0.1064	0.0304	1.1121	0.0105	0.0006
Radio	No	0.0819	0.5150	0.0001	-0.0113	0.0573	0.0000
	Yes	0.0015	0.5150	0.0051	0.6281	0.0573	0.0006
Overhead projector	No	0.0779	0.3372	0.0064	-0.1575	0.1540	0.0019
	Yes	0.0055	0.3372	0.0907	2.2489	0.1540	0.0276
Computer	No	0.0748	0.5498	0.0191	-0.4096	0.3323	0.0125
	Yes	0.0086	0.5498	0.1665	3.5701	0.3323	0.1093
Books other than text books	No	0.0209	0.7757	0.0621	-2.0254	0.6995	0.0859
	Yes	0.0624	0.7757	0.0208	0.6795	0.6995	0.0288

Source: Authors' calculations based on Young Lives School Survey 2011

A.9. Class facilities index by facility

Class facilities	No	Yes
Black board or white board	0.4870	-0.0030
Wall map	-0.6645	0.2888
Teacher's cabinet	-1.6670	0.2498
Teacher's desks	-4.6185	0.0454
Sufficient electric lights	-3.5707	0.1087
Electric fan	-1.6940	0.1659
TV	-0.1356	1.4547
Video player or DVD	-0.0030	0.4870
Radio	-0.0049	0.2751
Overhead projector	-0.0690	0.9848
Computer	-0.1794	1.5634
Books other than text books	-0.8869	0.2976

Source: Authors' calculations based on Young Lives School Survey 2011

A.10. Description of variables

	Obs.	Mean	SD	Min	Max
Math year-end test score	1,051	548.40	95.95	258.32	811.12
Vietnamese year-end test score	1,051	513.34	98.23	242.67	758.16
Attend FDS	1,051	0.43	0.50	0.00	1.00
Math test score	1,051	502.39	95.52	185.19	749.57
Vietnamese test score	1,051	498.71	96.30	217.85	727.02
Class facilities	1,051	-0.03	0.78	-5.86	1.78
Teacher's years of education	1,051	18.89	2.02	11.00	21.00
Lack of resource	1,051	0.50	0.50	0.00	1.00
Extra classes attendance	1,051	0.34	0.48	0.00	1.00
Social background	1,051	-0.01	0.83	-3.17	2.37
Boy	1,051	0.49	0.50	0.00	1.00
Minority	1,051	0.08	0.27	0.00	1.00
No sibling	1,051	0.33	0.50	0.00	1.00
1 or 2 siblings	1,051	0.47	0.50	0.00	1.00
More than 2 siblings	1,051	0.20	0.40	0.00	1.00
No older sibling	1,051	0.47	0.50	0.00	1.00
Urban areas	1,051	0.67	0.47	0.00	1.00
Age in month	1,896	119.63	3.69	106	136
Boy	1,896	0.49	0.50	0	1
Minority	1,896	0.15	0.35	0	1
No older sibling	1,896	0.46	0.50	0	1
Urban areas	1,896	0.20	0.40	0	1
Caregiver highest grade	1,896	6.41	3.53	0	12
Wealth index in 2006	1,896	0.44	0.18	0.01	0.85

Source: Authors' calculations based on Young Lives School Survey 2011

A.11. Selection equation estimation results

	Probability of being in Grade 5 in 2011
Age in month	0.161*** (0.00917)
Boy	0.0861 (0.0625)
Minority	-0.658*** (0.108)
No older sibling	-0.0339 (0.0630)
Urban areas	-0.191** (0.0871)
Caregiver highest grade	0.00154 (0.0116)
Wealth index in 2006	0.110 (0.243)
Constant	-19.04*** (1.096)
Observations	1,896

Source: Authors' calculations based on Young Lives School Survey 2011

A.12. Whole estimation results of learning progress in Mathematics (2SLS)

	(1)	(2)	(3)	(4)	(5)
	All	Q1	Q2	Q3	Q4
FDS attendance	-16.37*** (5.151)	-14.38 (11.62)	-25.33* (13.91)	-5.688 (56.77)	-19.98 (21.08)
Math test score	0.548*** (0.0259)	0.365*** (0.0531)	0.516*** (0.0712)	0.744** (0.308)	0.540*** (0.106)
Vietnamese test score	0.0653** (0.0265)	0.0828 (0.0519)	0.000170 (0.0738)	0.0335 (0.302)	0.162 (0.119)
Class facilities	9.833*** (3.197)	10.44** (4.934)	10.19 (9.117)	21.74 (45.31)	1.775 (17.67)
Teacher's years of education	-0.250 (1.170)	-1.054 (2.137)	-1.982 (3.489)	-0.286 (13.76)	1.596 (5.322)
Lack of resources	-5.552 (5.347)	-11.41 (10.35)	-5.773 (14.10)	-11.13 (63.18)	6.200 (24.60)
Extra-classes attendance	-1.981 (4.939)	-14.38 (10.39)	-4.082 (14.37)	-0.279 (55.49)	11.91 (20.13)
Social background index	8.826*** (3.198)	-9.405 (9.411)	-13.84 (52.60)	79.42 (191.2)	30.49 (23.13)
Boy	-2.053 (4.666)	-1.294 (9.947)	0.413 (13.33)	22.57 (71.06)	-5.812 (23.12)
Minority	5.772 (10.50)	-7.628 (14.28)	-13.63 (25.24)	29.55 (163.7)	-41.35 (165.9)
Has 1 or 2 siblings	16.85*** (5.311)	33.73*** (11.05)	23.52* (14.21)	3.015 (61.21)	13.35 (23.37)
Has more than 2 siblings	0.810 (6.565)	0.986 (12.48)	-0.696 (17.59)	0.230 (73.41)	3.175 (33.38)
Has older sibling	7.133 (4.677)	-4.183 (11.51)	2.602 (13.16)	7.206 (55.15)	4.440 (23.56)
Urban areas	-14.08** (5.998)	-2.922 (14.42)	-3.682 (15.90)	-13.57 (64.14)	-16.72 (26.08)
Inverted Mills Ratio	9.253 (9.592)	66.21 (89.17)	103.3 (118.6)	419.8 (1,001)	-162.9 (155.0)
Constant	249.1*** (28.84)	321.6*** (60.24)	326.8*** (80.62)	137.0 (326.8)	153.2 (131.6)
Observations	1,055	269	269	274	265

Source: Authors' calculations based on Young Lives School Survey 2011 and YLs R2 survey 2006.

A.13. Whole estimation results of learning progress in Vietnamese (2SLS)⁵⁶

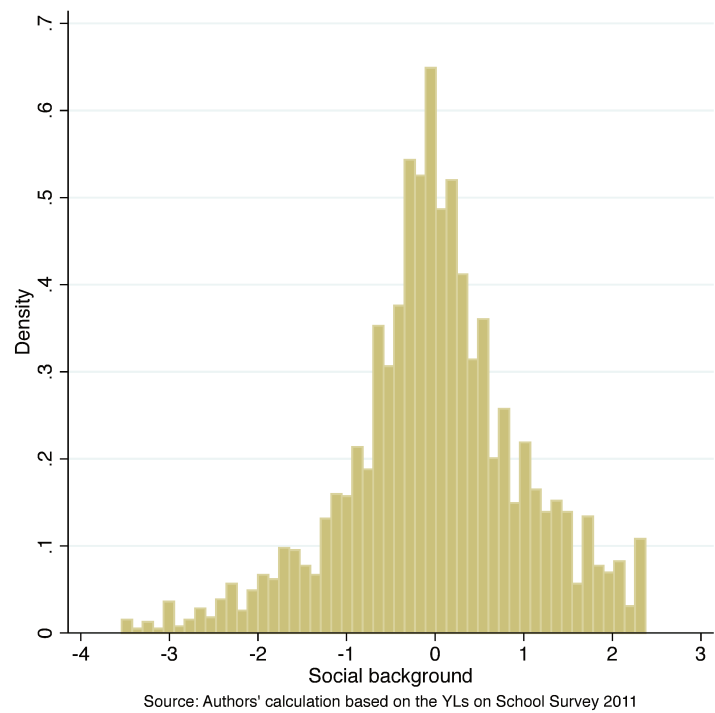
	(1) All	(2) Q1	(3) Q2	(4) Q3	(5) Q4
FDS attendance	-23.90*** (5.631)	-57.45** (26.96)	-19.27 (4.966e+08)	-5.498 (12.96)	-25.36 (30.99)
Math test score	0.0959*** (0.0280)	-0.0374 (0.120)	0.0882 (2.493e+06)	0.212*** (0.0706)	0.141 (0.155)
Vietnamese test score	0.452*** (0.0292)	0.353*** (0.122)	0.467 (2.595e+06)	0.365*** (0.0689)	0.611*** (0.173)
Class facilities	3.096 (3.411)	6.869 (11.08)	-0.644 (3.224e+08)	12.40 (10.33)	5.917 (25.87)
Teacher's years of education	0.245 (1.275)	2.394 (5.041)	-1.980 (1.240e+08)	-3.955 (3.147)	1.098 (7.629)
Lack of resources	-18.49*** (5.836)	-0.673 (23.73)	-22.40 (5.025e+08)	-39.30*** (14.34)	-15.51 (35.63)
Extra-classes attendance	3.630 (5.398)	-28.41 (24.39)	-11.09 (5.112e+08)	26.53** (12.66)	15.75 (29.11)
Social background index	7.535** (3.495)	-5.001 (22.12)	-15.76 (1.850e+09)	-85.79** (43.60)	-6.662 (33.46)
Boy	19.01*** (5.069)	2.214 (24.84)	22.79 (4.596e+08)	20.64 (12.88)	24.10 (32.00)
Minority	2.601 (11.31)	-30.67 (39.23)	-21.78 (8.671e+08)	-17.38 (36.39)	87.73 (257.1)
Has 1 or 2 siblings	8.267 (5.793)	2.340 (25.62)	7.876 (5.028e+08)	7.771 (14.04)	6.551 (33.49)
Has more than 2 siblings	-2.081 (7.169)	-3.538 (28.92)	-7.179 (6.221e+08)	8.235 (16.80)	-19.09 (48.42)
Has older sibling	-9.095* (5.091)	-4.091 (27.28)	0.250 (4.591e+08)	-7.001 (12.87)	-17.66 (44.84)
Urban areas	-16.84*** (6.515)	2.483 (33.18)	-11.42 (5.629e+08)	-41.92*** (14.37)	-19.53 (45.04)
Inverted Mills Ratio	3.394 (10.54)	-177.0 (199.8)	-3.588e+09 (4.372e+17)	94.89 (245.2)	-245.0 (702.5)
Constant	254.0*** (31.27)	333.4** (137.3)	287.8 (2.855e+09)	338.7*** (72.65)	171.5 (225.8)
Observations	1,053	269	269	273	265

Source: Authors' calculations based on Young Lives School Survey 2011 and YLs R2 survey 2006.

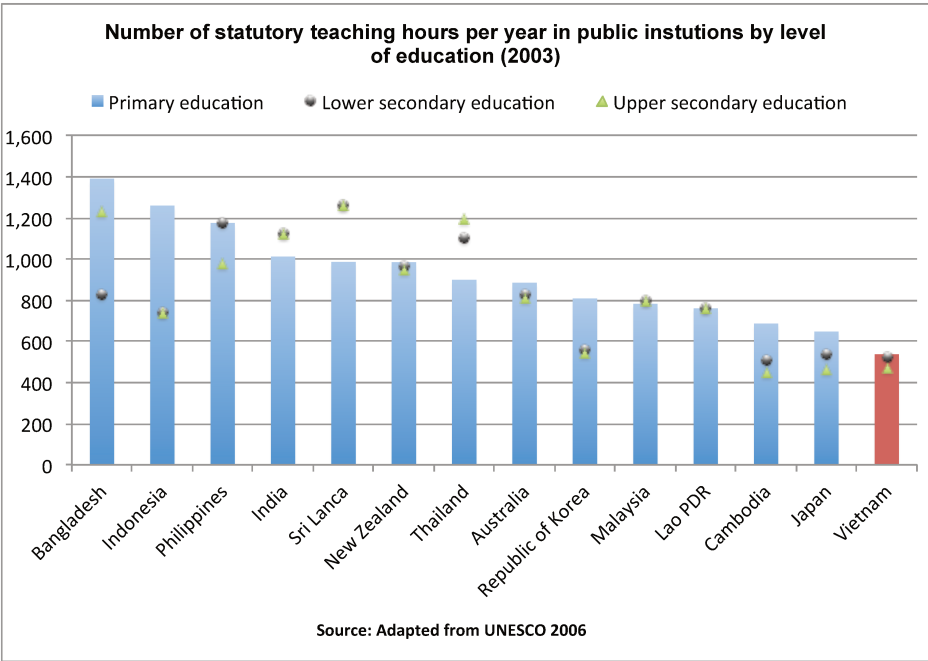
⁵⁶ For the Model (4), an OLS model was estimated for all YL children born in 2002 and from social background quartile 2 in Grade 5 in 2011 and consequently surveyed in the School Survey.

B. Figures

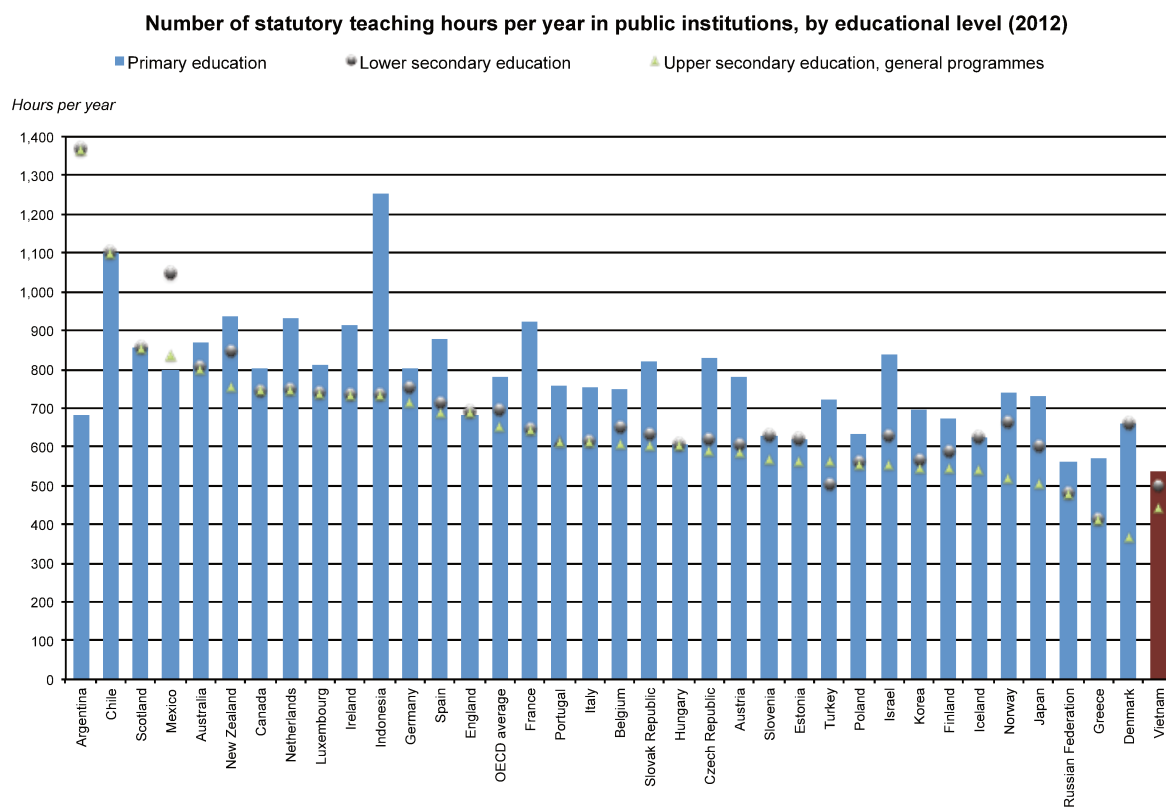
B.1. Distribution of students across the social background index



B.2. Statutory teaching hours, 2003



B.3. Statutory teaching hours, 2012



Source: Adapted from OECD (2014)

Chapter 3: The relative importance of skills in predicting labour market earnings in Vietnam

Abstract

While there is consensus on the predictive power of both cognitive and non-cognitive skills in explaining earnings or wages, the relative importance of these skills is still debatable. Contrary to abundant studies that provide mixed evidence on the effects of these skills for developed countries, little evidence exists for developing countries. This paper aims to contribute to the controversial evidence, particularly for the developing world. It uses the longitudinal data collected for the Old cohort born in 1993-1994 from the Young Lives in Vietnam, a developing country, to examine the relationship between both skills and labour outcomes. The paper applies the structural equation modelling, which allows latent cognitive and non-cognitive skills, together with individual, family and community characteristics, determine the observed measures of skills and the person's choices on education and labour outcomes. The regression results imply the higher effect of non-cognitive than cognitive ability on employment and earnings. Of the sample, the low-skilled are more likely to be employed, with a higher predictive power of the cognitive skill. Among the employment, those with higher cognitive skill tend to work as self-employed, while the higher non-cognitively tend to be wage workers and often have higher earnings.

1. Introduction

Both cognitive and non-cognitive skills are strong predictors of labour market outcomes (Almlund et al., 2011; Cunha & Heckman, 2009; Heckman, Stixrud, & Urzua, 2006). Cognitive skills are "identified with intelligence and the ability to solve abstract problems", and measured by the cognitive test scores, for instance, the IQ test. Non-cognitive skills are personality traits that weakly correlate with measures of intelligence. Examples of non-cognitive skills are self-esteem or self-efficacy. While there is consensus on the predictive power of both types of skills in explaining earnings or wages, the relative importance of these skills is still debatable. Profuse literature reveals the predominant role of cognitive skills on earnings (Almlund et al., 2011). There is also evidence showing that non-cognitive skills might be as important as or even more important than cognitive skills in predicting the labour market outcomes (Brunello & Schlotter, 2011). Mixed evidence derives from the use of different measures of non-cognitive skills and different methodologies in their studies.

Vietnam would be an interesting case study. On the one hand, Vietnam has been performing well in the education sector, with impressive results on student's acquisition of cognitive skills. Vietnam was ranked 12th out of 76 participating countries in the global ranking published by the Organization for Economic Cooperation and Development (OECD) in 2015 based on the 15-year-old students' performance in math, science and reading in the Program for International Student Assessment (PISA) in 2012. The rankings also put Vietnam above the OECD average as well as above the United States, Australia and the United Kingdom. The results from the STEP⁵⁷ surveys (Roseth et al., (2016) also show that Vietnam's labour force held the first in the literacy score and the second in the intensity of use in reading and writing skills among developing countries. On the other hand, the formal education and training system has provided the workforce with a low level of non-cognitive skills. Except for the highest score for Emotional Stability, other non-cognitive skills of the labour force (Extraversion, Openness, Agreeableness, Decision Making and Hostile Bias) are among the lowest one-third, especially the Conscientiousness and Grit stood the lowest second and third (ibid.). The report by the World Bank (2014) also claimed the shortfall of skills, inclusive of 'new' skills such as cognitive skills, critical thinking, teamwork and communication skills. Also, the labour productivity in Vietnam is among the lowest in the Asia-Pacific region,⁵⁸ in spite of significantly high labour productivity growth rated compared with other regional countries. Human capital characterised by skills of the labour force is one of the primary sources of labour productivity. Given the high level of cognitive skills and poor non-cognitive skills and low labour productivity, the low productivity might arise from the small return of cognitive skills and high return of non-cognitive on labour market earnings. This undefined situation correct motivates the analysis of the relative importance of these skills in determining earnings.

⁵⁷ The World Bank's Skills Towards Employability and Productivity – Skill Measurement Program that consists of a household survey and an employer survey. The program collected information on supply and demand for skills in urban areas in 12 developing countries, including Bolivia, Colombia, Ghana, Lao PDR, Ukraine, Vietnam, and the Yunnan Province in China in 2012 and Armenia, Azerbaijan, Georgia, Kenya, and Macedonia FYR in 2013.

⁵⁸ According to the report by MPI-GSO (2016), the productivity in Singapore, Malaysia, Thailand, Philippines and Indonesia was nearly 18, 7, 3 and 2 times the level in Vietnam respectively in 2013.

In the meantime, the current stagnating economic growth and the ageing population have forced the country to hinge on the higher productivity of the workforce, which requires the 'right' skills for its economic growth in the coming years (MPI-GSO, 2016; Vietnam Socio-economic Development Strategy for the period of 2011-2020). The proof for the relative importance of skills in determining labour outcomes is thus crucial for suggesting which type of skills should be enhanced more intensively to improve labour outcomes, inclusive of earning capacity or labour productivity. In reality, many Government's intervention programs have dealt with the skill gaps in various age ranges from the early childhood to schooling age and adulthood (see Almlund et al., 2011). The programs could provide remedial education or tailored tools to raise the cognitive and non-cognitive skills. The effects of these programs on both skills were positive at different effect sizes, for different groups, and some were insignificant. Pieces of evidence also reveal the improvement of non-cognitive skills and earnings through the programs primarily focusing on cognitive skills or health or poverty reduction.

Understanding the link between skills and labour outcomes is appealing for Vietnam as a developing country, for which such evidence is fewer than for developed countries. Among few such studies (Cunningham et al., 2016; Nordman et al., 2015; Sahn & Villa, 2015; Acosta, 2015; Krishnan & Krutikova, 2013; Díaz et al., 2013), none analyse the effects of skills on labour outcomes with the treatment for the endogeneity bias. The bias comes from the measurement errors in measuring skills (arising from the use of test scores as proxies for skills, while skills are unobserved, and test score are determined by skills and other factors, for instance, efforts) and the reverse causality between skills, education and labour outcomes. Scant evidence on the effect of skills on labour market outcomes is available for Vietnam. Almost all existing studies inspect how educational level or attainment affect labour outcomes (Gallup, 2002; Moock, Patrinos, & Venkataraman, 2003; Liu, 2004, 2005, 2008, Pham & Reilly, 2007, 2009; Imbert, 2013; Oostendorp & Doan, 2013; Phan & Coxhead, 2013; Sakellariou & Fang, 2014). More recent papers investigate the relationship between health and skills of both types (Dercon & Sánchez, 2011; Dercon & Sánchez, 2013; Dercon & Singh, 2013; Sanchez, 2013). Or they inspect the link between both skills and educational outcomes (Glewwe, Krutikova, & Rolleston, 2013; Glewwe, Qihui, & Katare, 2012).

For the facts as mentioned above, this paper, therefore, aims to assess the relationship between both cognitive and non-cognitive skills and labour market outcomes. Specifically, the paper will examine how both skills determine employment outcomes, and which type of skills is more important in predicting earnings. Data for analysis is from Young Lives Household Surveys for the young people born in 1994-95 in Vietnam, with information on their cognitive and non-cognitive skills and labour participation collected at age 18-19, and rich information on diversified live aspects of these young people, their caregivers and families over time. The study follows the approach used in Heckman et al. (2006), the structural equation modelling, to analyse the relationship between skills and labour outcomes. Accordingly, for each outcome of interest, a system of equations is specified. In the system, measures of latent cognitive and non-cognitive skill, together with individual, family and community characteristics, determine individual's observed measures of skills and their choices on education or one labour outcome simultaneously. The model specification addresses the measurement errors in skills and the reverse causality between skills and education. The reverse causality between skills and labour outcomes is expected to be unlikely in this paper because the employees seem not accumulate the skills from working or they had recently entered the labour market (only 9 percent of the current employed were doing paid jobs at their age of 15).

The contributions of this study are twofold. First of all, it is the first on the link between skills and labour market outcomes for Vietnam, with the value-added made on the literature gap in the effects of skills. Second, the study contributes to debatable evidence on the relative importance of cognitive and non-cognitive skills for developing countries and gives useful hints on which skill should be improved more during schooling for better labour outcomes.

Key findings of the paper imply the significant effects of cognitive and non-cognitive skills, with a higher effect of non-cognitive skill than cognitive skill on labour market outcomes. Young people aged 18-19 with low skills are more likely to work, and the predictive power of the cognitive skill is higher than the non-cognitive. Of the employed, the individuals with higher cognitive skill are more likely to work as self-employed, and the higher non-cognitively skilled are more likely to be wage worker and often have higher earnings.

The paper is structured as follows. Next section reviews the relevant literature on skills and the relationship between skills and labour outcomes. Section 4 introduces the concepts and measurement of skills. Section 3 describes the data source of the paper. Section 4 details the methodology for estimating the earnings return to skill. Section 5 gives the empirical results. Section 6 concludes the paper and discusses policy implications.

2. Skills and labour market earnings: a literature review

2.1. International literature

There is abundant literature on the relationship between cognitive and non-cognitive skills and employment outcomes. Most of the literature comes from the studies for developed countries, where longitudinal, nationally representative data sets on both types of skills and labour outcomes are available, such as the United States' National Longitudinal Survey of Youth (NLSY) or the German Socio-Economic Panel (GSOEP). These real datasets make the studies for developed countries golden standards regarding methodology employed, and enable the analysis of how skills cause labour outcomes in some cases.

Numerous shreds of evidences for *developed countries* showing that both cognitive and non-cognitive skills are strong predictors of labour market outcomes can be found in the literature review papers such as Hanushek & Woessmann (2008), Almlund et al. (2011), Hanushek (2014). However, the relative importance of skills in determining labour outcomes is still debatable. According to Almlund et al. (2011), in general, "Conscientiousness and Neuroticism are associated with job performance and wages to a similar but at a lesser degree than cognitive ability". The correlation between IQ and job performance is nearly 0.6 in one study. In other studies, Conscientiousness is the second associated with job performance, with the predictive power being a half of IQ. While the role of IQ increases with job complexity (for example, cognitive skills are more important for professors or managers than for unskilled workers), the role of Conscientiousness is invariant to task complexity. Conscientiousness, Emotional Stability (Neuroticism) facets, including locus of control⁵⁹ and self-esteem⁶⁰ determine

⁵⁹ one's belief that his/her life outcomes are the result of fate/luck or of his/her own actions

⁶⁰ one's subjective estimation of his/her own worth

choices on occupations and wages/earnings, at varying magnitudes across occupation and levels of education, but still at smaller levels than academic skills. Most recently, the findings from OECD (2015b) proves that the rise in cognitive skills rather than in non-cognitive skills raises income in most countries, except for Canada and United Kingdom.

While cognitive skills are the influential factor of earnings in many studies, a new growing literature emphasises the equally or even more important effects of non-cognitive skills on labour income. Several papers have corroborated the stronger role of non-cognitive over cognitive skills on the outcomes in the US and Europe (Almlund et al., 2011; Brunello & Schlotter, 2011), particularly the outcomes for young people (Gutman & Schoon, 2013). More specifically, the effect of non-cognitive skills is much stronger at the low end of the earnings distribution (Brunello & Schlotter, 2011) and for a lower level of job complexity or educational attainment (Almlund et al., 2011) in these countries.

Most of the existing evidence address the association rather than causality in making inference on the effect of skills on outcomes in labour market. This limitation is due to the methodological challenges, which are the problem of endogeneity and reverse causality. Endogeneity bias might arise from the measurement errors in the measures for skills, or the endogenous investment in skills. Measurement errors are likely if the task performance or test scores are proxies for skills, while the performance of tasks/tests depends on not only the skills but also on other factors such as efforts, incentives (more details in the subsequent section on the concept and measurement of skills). The interaction between investment and skills and labour outcomes (for instance, schooling improves skills and skills raise schooling, which then affects labour outcomes) would also cause biased estimated effects of skills. Besides, there might exist the reverse causality – the interplay between skills and the results (for example emotional stability might raise earnings, and higher wages might increase emotional stability) that may bias the estimated effects of skills. The risk of reverse causality probably arises in studies using cross-section data with skills and outcome observed simultaneously. There is a trade-off between this risk and the potential for bias estimation caused by using skills before entering the labour market in analysing the relationship between skills and labour outcomes. Because skills evolve over time, the skills in the past do not adequately capture the current skills, and the effect of skills on labour outcomes are thus distorted.

Many studies use the skills measured before individuals enter the job market as the treatment of the reverse causality. For example, both (Murnane et al., 2001) and Drago (2011) use NLSY 1979 to estimate Mincer earnings models for employees in 1987 (at age 27-28) with the inclusion of skills measured during high school years in 1980. The former study examines the effect of cognitive skills (measured by academic skills in math, reading and speed in doing an academic test) and non-cognitive skill (measured by self-esteem) on wage determination. Estimation results show the increase by nearly 10 and 4 % in wage for each standard deviation increase in academic skills and self-esteem respectively. The evidence of Drago (2011) using NLSY reveals that each standard deviation in self-esteem in 1987 is associated with from 4 percent increase in wage. The effect size of self-esteem doubles when the author corrects for endogeneity bias by instrumenting contemporaneous self-esteem by the self-esteem measured in 1980. Both studies only provide the evidence of an association between skills and wage/earnings, for they have not addressed the measurement errors in test scores, and have not taken into account the endogeneity of schooling or investment in estimating the effects

of skills. Also, the use of lagged skills as proxies for current skills in their analysis might generate bias results on the effects of skills.

Some studies give the evidence of the causal effect of skills on outcomes. Controlling for the measurement errors in measuring skills, reverse causality between skills and labour outcomes as well as between skills and schooling, Heckman et al. (2006) estimate the effects of different level of education on measured cognitive and non-cognitive skills. They find that education raises both types of skills, which in turn improve the labour outcomes. Skills affect earnings through its effects on education, then through effects of education on earnings. The effects of skills on earnings indirectly through the channel of education are evidenced in more recent studies. For instance, Heckman et al. (2011) in estimating sequential educational choice, prove that the personality (measured by participation in risky behaviour) in adolescence impacts earnings at the age of 30 through its effect on education. The direct effects of personality traits become vague given educational level. Cunha et al. (2010) estimate the technology of skill formation using longitudinal data on children with various measures of child skills and parental investment. This technology is a dynamic factor model to analyse the development of both cognitive and non-cognitive skills during childhood, controlling for endogenous investment in skills using shocks to family income and other instruments. Both Heckman et al. (2006) and Cunha et al. (2010) develop frameworks which are structural models to address the measurement errors in measuring skills. They relax traditional normality and linearity assumptions and do not rely on the measure of variances explained in estimating effects of skills on outcomes over the full set of relationships. They use the extension of factor analysis to represent proxies for latent factors (or latent skills). In Heckman et al. (2006), lifetime experiences and investment determine latent factors and other variables. In Cunha et al. (2010), the latent factors are allowed to determine investment and experiences. Their estimates of latent factors are corrected for “the effects of spurious feedback and separate proxies from factors”, and reflect changes over the life cycle in response to experience and investment (Almlund et al., 2011).

In contrast to the rich literature on the effect of skills labour outcomes for developed countries, none of the available studies for developing countries provides such evidence. This shortcoming is mainly due to there has been no data that allow for causal identification strategy. Either longitudinal data on individual cognitive and non-cognitive skills (from childhood to adulthood) for these countries is unavailable or unrepresentative or of small sample size.

Corroborating the evidence for developed countries, the related studies for **developing countries** all confirm that skills positively relate to wage or earnings. The studies include (Cunningham, Torrado, & Sarzosa, 2016; Nordman, Sarr, & Sharma, 2015; Sahn & Villa, 2015; Acosta, 2015; Krishnan & Krutikova, 2013; Díaz et al., 2013). The majority of the evidence shows the predominant role of cognitive skills over non-cognitive on labour outcomes. The measures of skills in their studies are rich in both cognitive skills, which could be reading, peabody receptive language, verbal fluency, working memory, numeracy, and non-cognitive, including Openness to experience, Conscientiousness, Extraversion, Agreeableness, Emotional Stability, Grit, Hostile, Decision making. Given various country contexts and methodologies applied to estimate earnings returns to skill, the estimated effect size and relative significance of skills differ across studies. None of the studies identifies the causal effect of skills on labour outcomes. Except for Nordman et al. (2015) and Krishnan & Krutikova (2013), all of

these studies control for the endogeneity of schooling in skills. However, only Cunningham et al. (2016) and Acosta et al. (2015) take into account the measurement errors in measuring skills.

Specifically, Díaz et al., (2013) and Sahn & Villa (2015) examine the effect of both cognitive and non-cognitive skills on earnings using both OLS and instrumental variable regressions to deal with the endogeneity of education and skills in estimating earnings. Díaz et al., (2013) use observed measures of skills to instrument schooling, and then use schooling residuals to instrument measured skills. Estimation results reveal that schooling has a higher effect than skills in Peruvian labour market. Conditional on skills, each additional year of school is associated with an increase of 15 percent in earnings. Both the overall cognitive skill and the perseverance facet of Grit has the equal effect size of 9% on earnings, conditional on schooling. Lower returns to earnings are observed for agreeableness (8%) emotional stability (5%). In the study of Sahn & Villa (2015), the grade attainment cognitive test scores are instrumented by personality traits and past household shocks. Besides, the authors also correct for the selection bias arising from sorting into the informal sector. The study shows the counterintuitive proof of negative earnings returns of cognitive skill, and the negative returns are robust across estimation models. Meanwhile, the estimates of the effect sizes of personality traits on earnings from the instrumental variable model are greater than from the OLS estimation. The evidence provided by Nordman et al. (2015), in explaining gender wage gap in the formal sector of Bangladesh using employer-employee data and both OLS and quantile regression, also give support to the stronger role of cognitive skills in determining wage. Non-cognitive skills prove to be significant in certain parts of the conditional wage distribution, especially for the female group. Krishnan & Krutikova (2013) employ non-randomized impact evaluation of a program to improve non-cognitive skills of children and adolescents in slums in Bombay, India. They found that the intervention raised self-efficacy and self-esteem by about one standard deviation, and life evaluation and aspiration by about half standard deviation. Their descriptive analysis shows that both self-efficacy and self-esteem positively correlate with the wage.

In addition to OLS with the inclusion of observed measures as proxies for skills, Cunningham et al. (2016) and Acosta et al. (2015) apply structural models to capture individual ability rather than measured skills like the remaining authors. Both authors use STEP - in Columbia for the former and in Peru for the latter – with the cross-section data on skills and labour outcomes. In their studies, skills are also latent variables specified by a system of equations on observed measures of skills and other factors determining skills. Results from Cunningham et al. (2016) emphasise the very robust role of verbal fluency and numeric ability on wage and employment choice in Peru. Meanwhile, except for Grit emerging strongly for most outcomes, the significance of other non-cognitive skills on wages or employment choice varies across estimation models. Acosta et al. (2015) documents that for Columbia, cognitive skills strongly and positively correlate with the probability to have a formal job or high-quality occupation and with earnings, while non-cognitive skills strongly relate to labour market participation. Although both Cunningham et al. (2016) and Acosta et al. (2015) address the measurement errors in skills, the reverse causality between skills and labour outcomes remains. There has been no study with causal identification of the effect of skills on employment outcomes for developing countries.

2.2. Related literature for Vietnam

There have been increasing number of studies on the cognitive and non-cognitive skills for Vietnam. Most of these studies rely on the data from the Young Lives Surveys⁶¹ in Vietnam, for its rich information on diversified live aspects of children and their caregivers and families, including cognitive and non-cognitive skills and labour outcomes, in 15 years. Only the World Bank's Vietnam Development Report 2014 (World Bank, 2014) uses a different source of data, the STEP⁶² 2011-2012. The report provides the profile of skills of the workforce and demand for skills in urban areas in Vietnam, without any section on how the skill gaps determine labour market outcomes.

The studies using Young Lives data have focused on the skill formation (Borga, 2015; Sanchez, 2013; Dercon & Singh, 2013; Rolleston & James, 2011), especially the role of early nutrition in forming skills (Georgiadis et al., 2016; Le & Tran, 2015; Dercon & Sánchez, 2013; Dercon & Sanchéz, 2011). Fewer studies look in the link between skills and educational outcomes (Singh, 2014; Glewwe et al., 2013; Glewwe et al., 2012) rather than labour outcomes. These researches mainly make the focus on cognitive rather than non-cognitive skills.

Of the available studies on non-cognitive skills for Vietnam, there is only one that looks at the correlation between non-cognitive skills and learning achievement – the Rolleston et al., (2013). According to the report, only academic confidence has a beneficial effect on learning progress, independent of academic performance. The two remaining skills, which are Academic Effort and Academic Stress, are insignificantly related to learning achievement.

Abundant literature on the role of education in determining wage growth and wage inequality is available for Vietnam. These studies mainly come with the availability of data from the VHLSSs - Vietnam Household Living Standard Surveys⁶³ (Gallup, 2002; Moock, Patrinos, & Venkataraman, 2003; Liu, 2004, 2005, 2008; Pham & Reilly, 2007, 2009; Imbert, 2013; Oostendorp & Doan, 2013; Phan & Coxhead, 2013; Sakellariou & Fang, 2014; Nguyen & Tran, 2016),. They show that higher level of education improved an individual's probability to be in employment and to earn more. There has been only paper investigating the earnings returns to skills in Vietnam, which is Sakellariou (2009). In this study, the returns to the cognitive skills – the skills on computer and foreign languages of university graduates in Vietnam, were estimated using the Heckman model with endogenous explanatory variables. The data employed is from the Higher Education Graduate Tracer Survey conducted in 1996 by the Ministry of Labour, Invalids and Social Affairs as part of the World Bank – Vietnam Education Financing Sector Study. Estimation results show the strongly positive contribution of these skills to

⁶¹ Young Lives is a longitudinal study on children's poverty based on household surveys of 12,000 children conducted in Ethiopia, India, Peru and Vietnam (see Section 6 in the General Introduction for more details)

⁶² The Skills Toward Employment and Productivity consists of a household and employer survey to collect information on the supply and demand for skills in Ho Chi Minh City and Hanoi in 2011-2012.

⁶³ Nationally representative survey to collect data on household income and expenditure, inclusive of employment and wage data, enabling the measurement of living standard and relevant indicators. The first survey was conducted in 1993, the second in 1998, then every two years from 2002 till now.

earnings premiums. No research on the link between both types of skills and labour earnings using the Young Lives data or STEP has been conducted yet.

3. Data:

3.1. Measures of cognitive and non-cognitive skills

The Young Lives in Vietnam gives data to measure a variety of cognitive and non-cognitive skills of children at different ages. At age 19, the cognitive skill is measured by the test score in Math and Language (in Vietnamese). The non-cognitive skill is measured by Self-efficacy, Self-esteem, Parent relations and Peer relations. (for definitions, statements and computation of the measures, see Section 6.2 of the General Introduction). Appendix C.1 gives the distributions of composite and individual skill indexes.

There is a high correlation between different skills of one type of skills (among cognitive skills or non-cognitive skills). Insignificant to weak correlation exists between individual skills across types of skills. Factor analysis of individual skills is conducted to find whether there are factors that represent cognitive skill and non-cognitive skill. The results of factor analysis cognitive and non-cognitive skills are provided in Appendix B.1. The pairwise correlation between skills shows the proper representativeness of the composite skill index for individual indexes. The associations between different indexes are much significantly stronger for cognitive than non-cognitive skill. Among the non-cognitive indexes, self-esteem and self-efficacy, and self-esteem and peer relations are observed with the highest correlation. Across cognitive and non-cognitive skills, the correlation is weak and observed for only some indexes. Of which, the self-efficacy positively correlates with Math, while Parent and Peer relations negatively correlate with cognitive composite skill and Math. It is noteworthy that highest level of education completed is highly associated with the composite and individual cognitive skills.

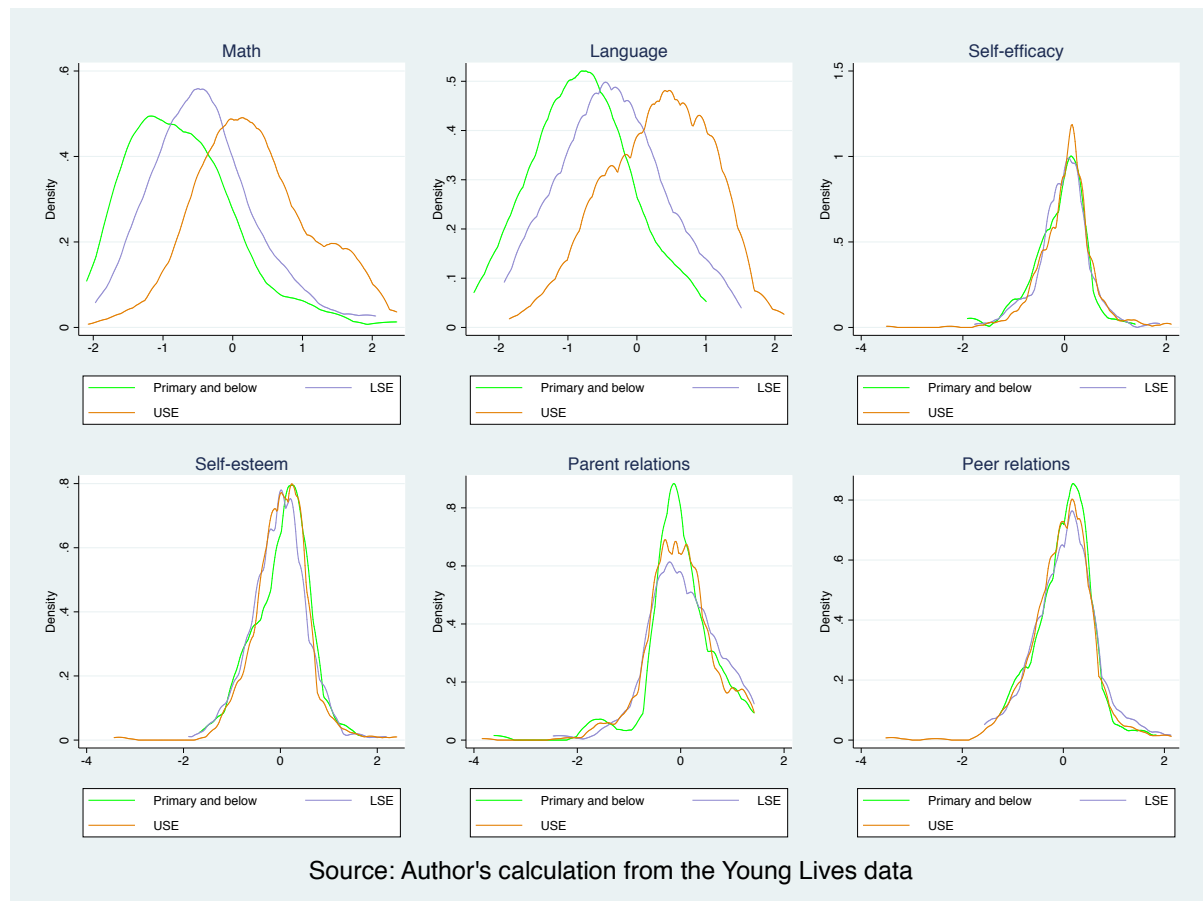
Table 3- 1: Pairwise correlation between skills (age 19)

	Cognitive skill	Math test score	Language test score	Non-cognitive skill	Self-efficacy	Self-esteem	Parent relations	Peer relation	Educational level
Cognitive skill	1								
Math test score	0.8981*	1							
Language test score	0.8982*	0.6133*	1						
Non-cognitive skill	-0.0287	-0.0156	-0.0341	1					
Self-efficacy	0.0558	0.0958*	0.00720	0.7462*	1				
Self-esteem	0.0407	0.0679	0.00710	0.8359*	0.5682*	1			
Parent relations	-0.1081*	-0.1353*	-0.0603	0.6053*	0.2363*	0.3203*	1		
Peer relation	-0.0936*	-0.1019*	-0.0643	0.7845*	0.3897*	0.5303*	0.3898*	1	
Educational level	0.4790*	0.4264*	0.4406*	0.00460	0.0786*	-0.0007	-0.0332	-0.0347	1

Figure 3- 1 suggests that schooling facilitates the students' acquirement of cognitive skills rather than non-cognitive skills. A Higher level of education comes together with higher skill levels of cognitive skills. For the non-cognitive skills, except for a larger share of those completed upper

secondary level with around the average level of self-efficacy and a higher proportion of the primary school with average parent and peer relations, the trend of skills across educational levels is unclear. It might be the jobs that have improved the non-cognitive skills of those who dropped out of school to work.

Figure 3- 1: Distribution of skill indexes by level of education

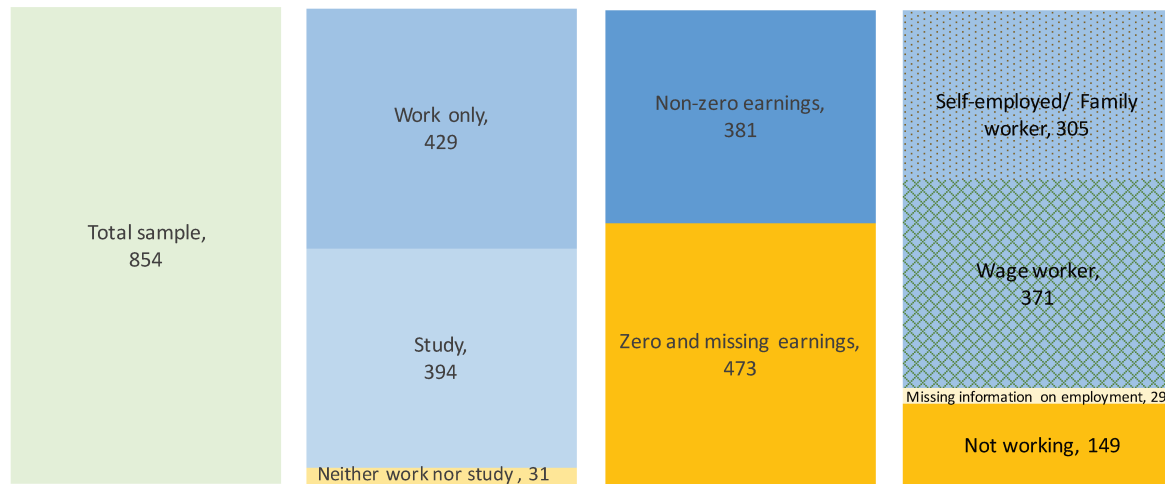


3.2. Data on employment and earnings

The panel sample for the Older Cohort covers 886 young people aged around 18-19. About 79 percent conducted paid or unpaid work(s). The information on employment and earnings are collected for the first (most important) paid and unpaid work activity only. The panel with non-missing data on labour market participation or earnings, which is the data for this study, includes 854 observations. The structure of the panel is provided in Figure 3- 2. Of the total sample, a half stopped their schooling and were working. Another 46 percent were going to school, either working simultaneously or not, and the rest neither work nor study. Among those having work activities, 55 percent are wage workers, and the remaining are self-employed or family workers. Less than a half of the sample have non-zero earnings (381 observations). The non-zero earnings are mainly observed for the wage workers (nearly 93 percent), while only 12 percent of the self-employed/family workers reported non-zero earnings. Therefore, it is possible to look at young people's choices on whether to enter the

labour market or not, then if work, on whether to work as wage workers or self-employed/family worker. But the analysis of the log of hourly earnings will be conducted for firstly those with non-zero earnings only.

Figure 3- 2: Sample structure



Source: Author's calculation from Young Lives data

Table 3- 2: Characteristics of the sample

	Obs.	Mean	SD
Male	854	0.48	0.50
Married/Cohabitation	854	0.11	0.31
Below primary education	832	0.02	0.15
Primary education	832	0.12	0.33
Lower secondary education	832	0.24	0.43
Upper secondary education	832	0.58	0.49
Bachelor or higher education	832	0.03	0.17
Mother completed below primary education	847	0.26	0.44
Mother completed primary education	847	0.31	0.46
Mother completed lower secondary education	847	0.28	0.45
Mother completed upper secondary or higher education	847	0.15	0.36
Father completed below primary education	835	0.20	0.40
Father completed primary education	835	0.28	0.45
Father completed lower secondary education	835	0.32	0.47
Father completed upper secondary or higher education	835	0.20	0.40
Household size	854	4.20	1.49
Household's shock(s)	854	0.41	0.49
Urban community	820	0.22	0.42

Source: Author's calculations based on Young Lives Surveys

Table 3- 2 shows that the sample is balanced in term of gender. Most of the young people had completed upper secondary education, or below, of which the upper secondary level accounts for 58 percent, and less than lower secondary 40 percent. At the time of the survey, 11 percent of the young people were married or cohabiting with their partners. On average, they were living in a four-member household, and mostly in rural areas. About 41 percent of these households had encountered a shock or some shocks since 2009. The shocks could be income related shocks (the death of livestock, crop failure, the loss of job) or illness or death of one of their family member. Their parents had less than upper secondary education (85 percent of mother and 80 percent of the father with less than upper secondary level).

4. Estimating returns to skills: empirical strategy

This paper estimates the effects of both cognitive and non-cognitive skills on young people's labour outcomes, including the probability of being employed, being wage-worker, or self-employed, and the log of hourly earnings. The conventional approach to the investigation involves the OLS (for earnings) and logit (for the labour market participation) estimation of equations. In the model, a labour outcome y of person i is specified as a function of skills, including cognitive (C) and non-cognitive (N) type, and a set of other observed factors that impact labour outcomes (x_i), and an error term ε_i that reflects unobserved factors and measurement errors. The error term is assumed to be independent of C_i , N_i and x_i variables:

$$y_i = \beta_C C_i + \beta_N N_i + \beta_x x_i + \varepsilon_i \quad (1)$$

The expected labour outcomes vary according to the β_C and β_N , the so-called returns to cognitive and non-cognitive skills. Using the conventional approach, the skill included in the equation can be a set of observed measures of individual skills (individual skill indexes) or a set of composite measures of cognitive and non-cognitive skill (cognitive and non-cognitive skill index). The inclusion of all individual skill indexes in the equation gives rise to the problem of simultaneity, for there is a high correlation between skills of one type and between two types of skills. This problem can be mitigated by using composite indexes, which are inferred using variable reduction techniques (simple average of normalised individual indexes, factor analysis or principal component analysis). However, because skills are not the only factors that affect the observed test scores (see Section 3: Skills: concepts and measurement for a detailed discussion), these skill indexes are not good proxies for skills. These problems cause biased and inconsistent estimates of the effects of skill this conventional approach.

The alternative way following the approach of Heckman et al., 2006 to deal with the measurement errors in measuring skills is assuming that skills are latent factors θ instead of observed skills as follows:

$$y_i = \beta_C \theta_i^C + \beta_N \theta_i^N + \beta_x x_i + \varepsilon_i \quad (2)$$

where θ^C, θ^N are estimated from observed measures of skills T and other observed factors x_i^C and x_i^N that impact the observed measures:

$$T_i^C = \alpha_C \theta_i^C + \alpha_x^C x_i^C + \epsilon_i^C \quad (3a)$$

$$T_i^N = \alpha_N \theta_i^N + \alpha_x^N x_i^N + \epsilon_i^N \quad (3b)$$

where T^C is a 2x1 vector of observed measures of cognitive skill (math and language test scores); T^N is a 4x1 vector of observed measures of non-cognitive skill (self-efficacy, self-esteem, parent relations and peer relations); all the error terms $\epsilon_i^C, \epsilon_i^N$ are mutually independent and independent of θ^C, θ^N, x and ϵ_i . Specifically, the equations (2), (3a) and (3b) together form the following structural model based on which the effects of skills are estimated (with all observed measured of skills in the order mentioned above):

$$\begin{cases} T_i^{Math} = \alpha_{Math}\theta_i^C + \alpha_x^{Math}x_i + \epsilon_i^{Math} \\ T_i^{Lang} = \alpha_{Lang}\theta_i^C + \alpha_x^{Lang}x_i + \epsilon_i^{Lang} \\ T_i^{Sef} = \alpha_{Sef}\theta_i^N + \alpha_x^{Sef}x_i + \epsilon_i^{Sef} \\ T_i^{Ses} = \alpha_{Ses}\theta_i^N + \alpha_x^{Ses}x_i + \epsilon_i^{Ses} \\ T_i^{Par} = \alpha_{Par}\theta_i^N + \alpha_x^{Par}x_i + \epsilon_i^{Par} \\ T_i^{Peer} = \alpha_{Peer}\theta_i^N + \alpha_x^{Peer}x_i + \epsilon_i^{Peer} \\ y_i = \beta_C\theta_i^C + \beta_N\theta_i^N + \beta_x x_i + \epsilon_i \end{cases} \quad (4)$$

Similar to (4), the effects of skills on educational outcome are estimated using the structural model (5), in which individual's choice on educational level E depends on skills and other observed factors x^E , and unobserved factors v_i that impact the choice as follows:

$$\begin{cases} T_i^{Math} = \alpha_{Math}\theta_i^C + \alpha_x^{Math}x_i + \epsilon_i^{Math} \\ T_i^{Lang} = \alpha_{Lang}\theta_i^C + \alpha_x^{Lang}x_i + \epsilon_i^{Lang} \\ T_i^{Sef} = \alpha_{Sef}\theta_i^N + \alpha_x^{Sef}x_i + \epsilon_i^{Sef} \\ T_i^{Ses} = \alpha_{Ses}\theta_i^N + \alpha_x^{Ses}x_i + \epsilon_i^{Ses} \\ T_i^{Par} = \alpha_{Par}\theta_i^N + \alpha_x^{Par}x_i + \epsilon_i^{Par} \\ T_i^{Peer} = \alpha_{Peer}\theta_i^N + \alpha_x^{Peer}x_i + \epsilon_i^{Peer} \\ E_i = \gamma_C\theta_i^C + \gamma_N\theta_i^N + \gamma_x x_i^E + v_i \end{cases} \quad (5)$$

The cognitive skills included only in (3a), and non-cognitive skill included only in (3b) rely on the assumption that only one latent cognitive factor and one latent non-cognitive factor mainly determines the individual cognitive and non-cognitive skill indexes respectively. The assumption is justified by the previous results of the factor analysis of individual skills, which show there exist factors representing cognitive and non-cognitive skill (see the previous section on measures of skills using Young Lives data). In this two-factor model (cognitive and non-cognitive factor in 3a and 3b), one coefficient of α_C and one of α_N must be normalized to unity, and the remaining coefficients are explained in proportion to the normalised coefficients. The identification of (3a) and (3b) is satisfied because the number of available skill indexes is six, including math, language, self-efficacy, self-esteem, parent relations, peer relations. The number is greater than the required minimum number of five test scores/indexes as stated by (Carneiro, Hansen, & Heckman, 2003): $n = 2f + 1$ (n is the number of scores; f is the number of factors).

Controlling for x , dependence across the measurements, choices and outcomes based on θ^C, θ^N in the model helps to deal with the measurement errors in skills, without taking into account

education in the labour outcome equations (as in traditional OLS and instrumental variable models)⁶⁴. With labour outcomes specified as solutions to maximization problems in response to constraint (inclusive of skills) and preferences, the estimation of the structural model provides real effects of skills on labour outcomes.

Data on skills observed at age 19 is used for the estimation of the models because they fully reflect the level of skills at the time of reference, while the skill observed in the previous round misses the part of skills accumulated between age 15 and 19. Additionally, following the survey design, the non-cognitive skill at age 15 captures narrower domain of the skill and are less reliable than the current skills (see Section 6.2 of the General Introduction). The reverse causality between skills and labour outcomes due to measures of skills observed contemporaneously with labour outcomes is expected to be unlikely in this paper. There was only a small share (9 percent) of the current employed worked when they were 15. Thus, their current skills seem not to be accumulated from working, or the causal effect of working/labour outcomes on the development of skills is negligible.

The structural models in this paper are estimated using the Stata command GSEM. The latent cognitive and non-cognitive skills are inferred from (3), in which the math test score and the self-efficacy index are fixed to be 1 (for the model identification as mentioned previously). The model imposes no assumption on the distribution of the latent factors and error terms in the equations on employment types. The remaining error terms follow normal distributions. Choices on educational and labour outcomes are examined by estimating (5) and (4) respectively.

The labour outcomes measured at the age of 18-19, comprising the probability of being employed, wage worker or self-employed and \ln of hourly earnings. Educational outcome at the age of 18-19 could be one of three levels, including primary and below, lower secondary, and upper secondary level and above. Educational choices and employment types are estimated using multinomial logit regressions, and the probability to be employed is estimated using logit regressions. Each of the outcomes is regressed on cognitive and non-cognitive skill, and a set of control variables (details in Appendix B.2).

Individual's decision to enrol or obtain a level of education is assumed to be affected by the latent skills, the local wage (measured by the community's average daily wage) and the local unemployment rate at this age affect individual's decision on whether to enrol and obtain. In other words, they weigh the benefit and cost associated with the level of education that they determine to pick. Other

⁶⁴ In the OLS, the inclusion of education in a labour outcome equation gives rise to the problem of endogeneity and reverse causality between education and skills (education raises skills, and skills affect education). Yet if education is excluded from the equation, the indirect effects of skills on labour outcome through education might not be fully captured, causing biased estimates of effect of skills. Applying instruments for education is one option but the method does not always ensure addressing the endogeneity of education because (i) good instruments are often debatable; (ii) the coefficients are estimated with large standard errors, (iii) the difference between OLS and IV estimates is small if sampling variations is controlled for; and (iv) if the returns to skills are heterogeneous, the IV estimates do not reflect the effect of education, but the returns for individual going to school as a result of changes in the values of the instrument (Heckman et al. , 2006).

personal, family and community characteristics are also allowed to impact the choices on educational levels. These factors comprise the individual's gender, the number of siblings, family's economic condition, parents' education, family's shocks, residential region (urban or rural). Family's characteristics were measured in the previous rounds of the survey. Parents' education refers to the highest educational level completed by the individual's father and mother (one of the five levels including below primary, primary, lower secondary, upper secondary, bachelor or higher level) at the individual's age of 14-15. Family's economic condition is represented by the household's wealth index when the person was 11-12 years old. Family's shocks refer to whether or not the household had ever encountered a shock during the period between the individual's age of 14-15 and 18-19.

For the labour outcomes including labour market participation, employment type and earnings, given the fixed cognitive and non-cognitive skills, the community unemployment rate and residential region, it is assumed that the family background (for instance parents' educational level) does not impact the earnings.

5. Skills and labour outcomes

5.1. Skills and labour earnings: descriptive analysis using the Young Lives data

Table 3- 3: Skills and earnings, by group

	Whole sample	Female	Male		Not working	Working	Wage worker	Self-employed	
Math	0.01	-0.02	0.03		0.19	-0.03 ***	-0.05	-0.03	
Language	-0.00	0.11	-0.12 ***		0.22	-0.05 ***	-0.10	-0.01	
Self-efficacy	-0.00	-0.06	0.06 ***		0.05	-0.01	-0.00	-0.02	
Self-esteem	0.00	-0.01	0.02		0.05	-0.01	0.01	-0.02	
Parent relation	-0.00	0.03	-0.03		0.08	-0.02	-0.05	0.04	*
Peer relation	0.00	-0.01	0.01		0.08	-0.01 *	0.01	-0.03	
Primary education	0.15	0.13	0.16		0.07	0.16 ***	0.19	0.14	*
Lower secondary education	0.24	0.21	0.28 **		0.10	0.27 ***	0.28	0.27	
Upper secondary education	0.61	0.66	0.56 ***		0.83	0.56 ***	0.53	0.59	*
Years of schooling	10.65	10.87	10.43 ***		11.65	10.44 ***	10.21	10.62	**
Monthly earnings (1,000 VND)	21,723	20,782	22,623		.	21,723	38,456	1,953	***
Hourly earnings (VND)	17,793	11,415	23,906		.	17,793	30,094	3,382	**
Number of observations	854	440	414		149	705	371	305	

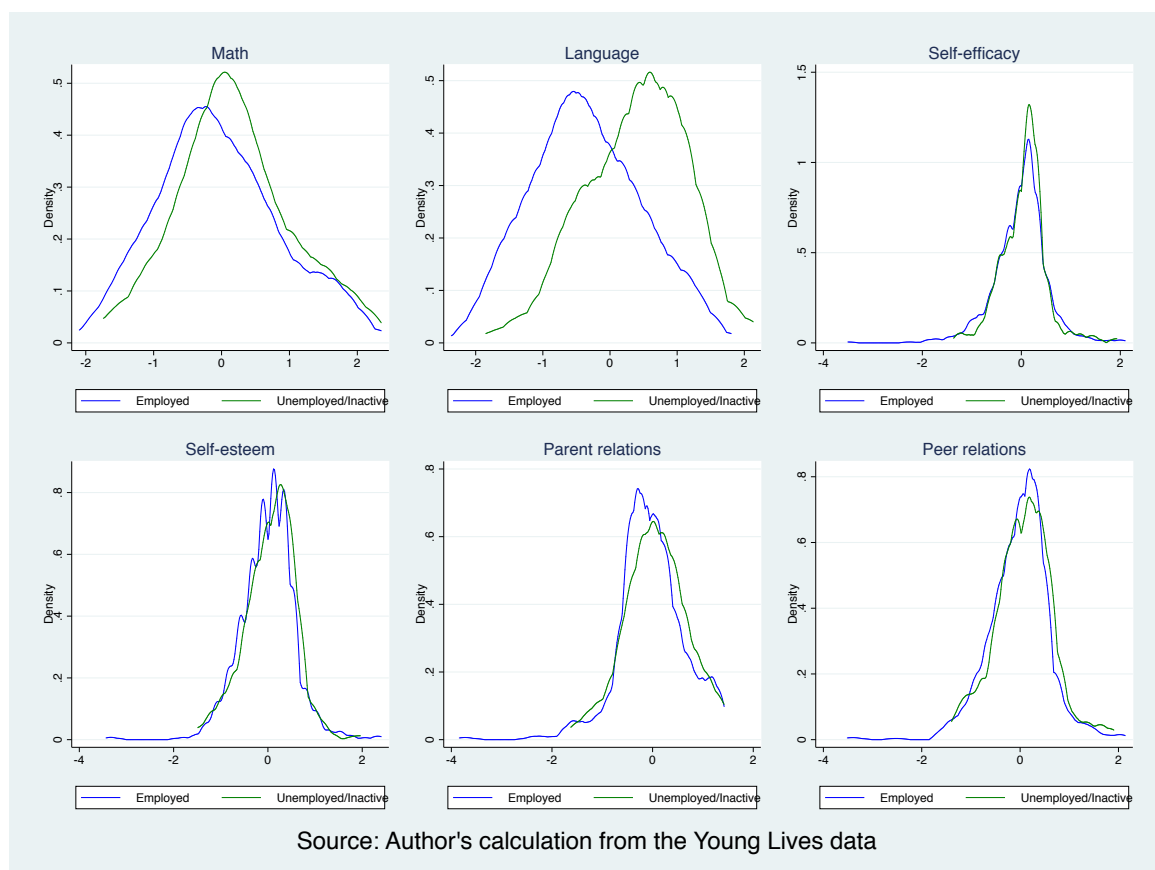
Note: difference significant at 1% (***), 5% (**), 10%(*)

Source: Author's calculation from the Young Lives data

Of the whole sample, the young women averagely were more educated than men. The proportion of women having at least upper secondary education 66 percent while the men's only 54 percent. Notably, the language skill is much better for women compared with men, but women had much lower self-efficacy than men. Men's hourly earnings double women. The higher earning capacity of men over women might derive from gender discrimination or gender gaps in skills.

The skills and earnings of young people also vary according to the choices⁶⁵ they made on labour market participation and employment types as shown in Table 3- 3. Three-quarters of the young people not working were students with the highest level of education, while nearly two third of the working youth stopped their schooling. Thus, those not working were more educated than the employed: on average 12 years of schooling and about 83 percent completed at least upper secondary education (versus ten years of schooling and 56 percent of the working group that have completed at least upper secondary education). The not working group achieved higher cognitive skills and often had a much better relationship with their peers. Figure 3- 3 illustrates these trends and additionally shows that greater share of the employed obtained the mean self-efficacy while the more of the not working youth had mean self-esteem and parent relation indexes.

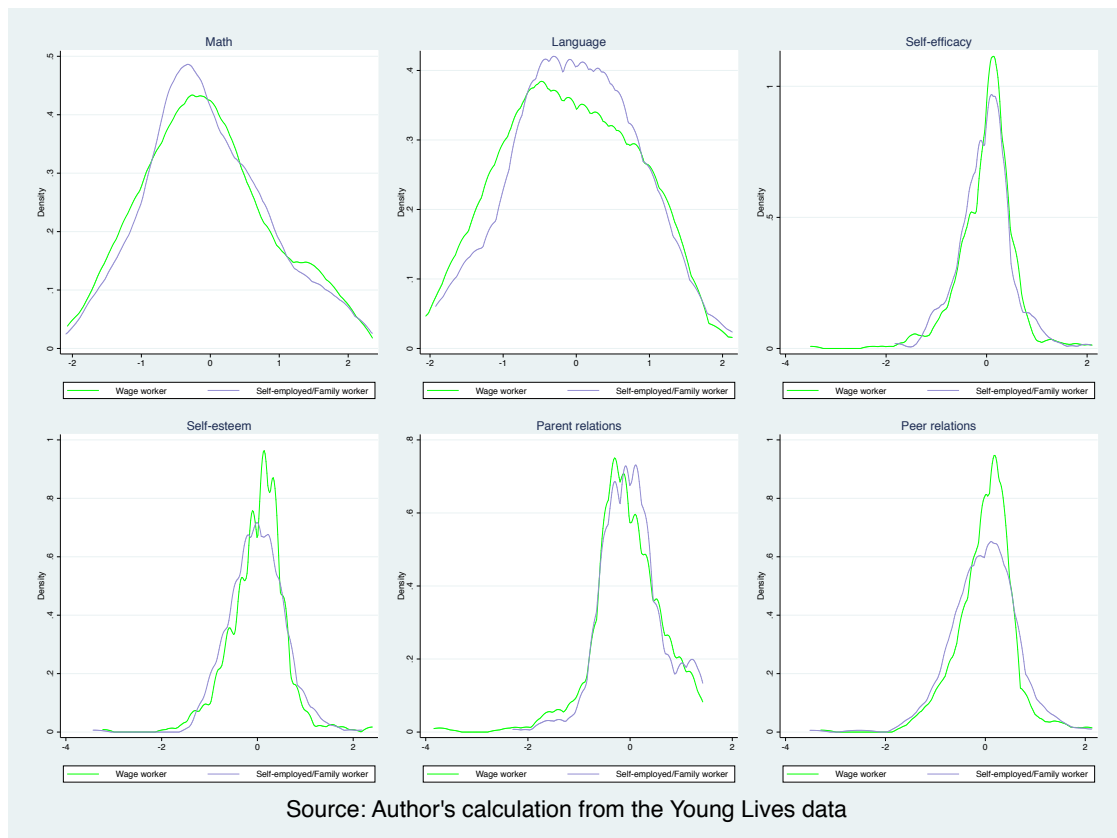
Figure 3- 3: Distribution of skills, by choice on labour market participation



⁶⁵ The group that neither work nor study are skipped due to few observations (31). Moreover, this group belongs to of the unemployed/inactive (not working) in the sample disaggregation by employment type below.

Differences in educational levels between young wage workers and self-employed/family worker is small. Better relationship with parents was observed for young self-employed than for the wage workers. Gaps in the average indexes of the remaining skills between two employment types are insignificant, but the skill distributions show that higher share of the self-employed obtained mean cognitive skills and a greater proportion of the wage workers have mean self-efficacy, self-esteem and peer relations indexes (Figure 3- 4). The gap in earnings might be not meaningful due to the feature of unrecorded income of the self-employed or family workers.

Figure 3- 4: Distribution of skills, by employment type



5.2. Skills and Education: estimation results

Estimation of the structural models shows that the latent cognitive and non-cognitive skills are significant in most equations on young people's outcomes. The results of tests for normality (in Appendix B.3) show that both estimated cognitive and non-cognitive skills are non-normally distributed. Thus, the assumption of the normal distribution of latent skills in the model will yield biased estimates of the effects of skills on youth's outcomes.

Estimation results of the skill indexes in Table 3- 4 suggest that the latent cognitive skill is more highly associated with the language skill, while latent non-cognitive skill more highly relates to individual's self-esteem and peer relations. People from economically advantaged family or more educated mothers have high individual cognitive skills. A higher level of mother's education is also observed for the individuals with higher self-efficacy. However, a better off family is associated with

lower self-esteem and peer relations. Young people in urban residence are more likely to obtain less math skill and parent relations.

Table 3- 4: Model of cognitive and non-cognitive skills

	Math	Language	Self- efficacy	Self- esteem	Parent relations	Peer relations
Cognitive skill	1.000 .	1.023*** (0.000)				
Non-cognitive skill			1.000 .	1.347*** (0.000)	0.819*** (0.000)	1.139*** (0.000)
Male	0.050 (0.405)	-0.237*** (0.000)	0.113*** (0.002)	0.038 (0.349)	-0.051 (0.272)	0.004 (0.920)
Number of siblings	-0.051* (0.054)	-0.089*** (0.001)	-0.016 (0.315)	0.002 (0.925)	-0.010 (0.615)	-0.017 (0.350)
Wealth index	1.410*** (0.000)	0.951*** (0.000)	0.012 (0.931)	-0.302** (0.042)	0.059 (0.726)	-0.459*** (0.004)
Mother's education	0.168*** (0.000)	0.109*** (0.003)	0.058** (0.014)	0.033 (0.201)	-0.030 (0.303)	0.026 (0.329)
Father's education	0.055 (0.122)	0.045 (0.205)	-0.015 (0.509)	0.027 (0.261)	-0.024 (0.393)	-0.012 (0.638)
Household's shocks	-0.047 (0.440)	-0.079 (0.184)	-0.033 (0.380)	-0.065 (0.111)	-0.042 (0.365)	-0.025 (0.569)
Urban community	-0.235*** (0.003)	0.006 (0.939)	0.020 (0.690)	-0.036 (0.510)	-0.138** (0.027)	0.067 (0.251)
Constant	-1.134*** (0.000)	-0.576*** (0.000)	-0.116 (0.168)	0.027 (0.768)	0.183* (0.078)	0.231** (0.017)

Note: p-value in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Source: Authors' calculations based on the Young Lives data

The densities of estimated latent cognitive and non-cognitive skills across educational levels displayed in Figure 3- 5 corroborates their non-normal distribution. As also confirmed by the estimation results of education equations (Table 3- 5), the people with higher cognitive skill often have a greater level of education. This trend is similar to the one in Figure 3- 1 on the distribution of individual skill indexes by educational attainment. The distribution of latent non-cognitive skill between educational levels has a clearer shape than that of the individual skill indexes. Accordingly, young people with upper secondary education are the most non-cognitively able, and the least is the group that completed lower secondary school. It appears that educational level does not significantly and positively relate to non-cognitive skills. People that completed primary level of education seem to have improved their non-cognitive skills outside of schools, probably at working places.

Table 3- 5 shows that family background is significant to individual's choice on level of education. Parents with better economic status, higher educational level, and fewer children (thus more care dedicated to their children) are more likely to have more highly educated children. Mother has a more influential role in children's educational choice than the father.

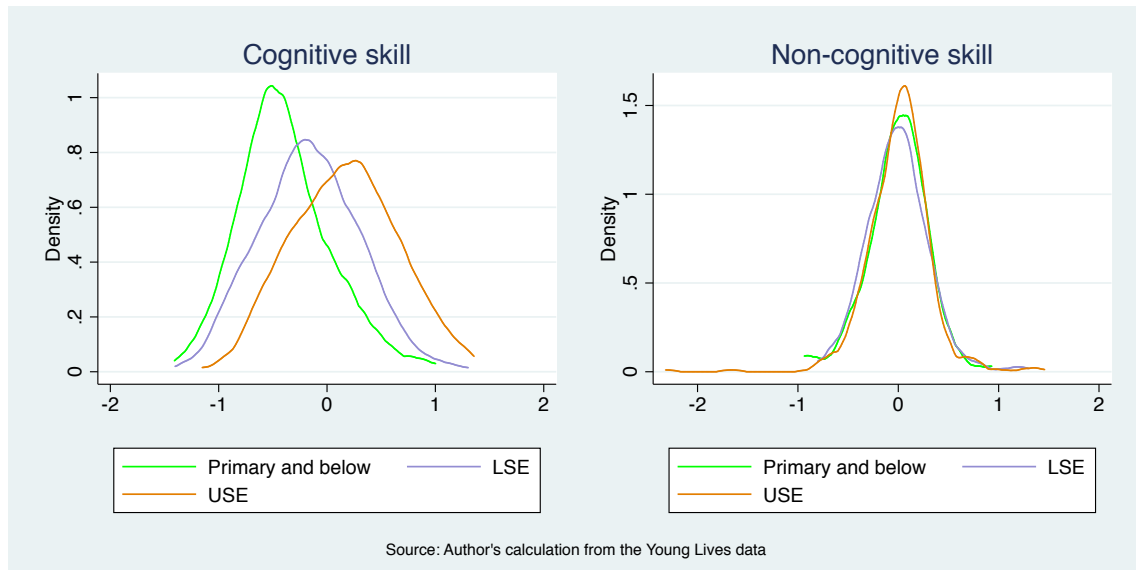
Table 3- 5: Model of cognitive and non-cognitive skills: Choices on educational level

	Lower secondary education	Upper secondary education
Cognitive skill	2.274*** (0.000)	4.197*** (0.000)
Non-cognitive skill	0.355 (0.532)	0.596 (0.338)
Male	-0.050 (0.895)	-0.748* (0.084)
Number of siblings	-0.623*** (0.000)	-0.636*** (0.000)
Wealth index	5.315*** (0.001)	9.926*** (0.000)
Mother's educational level	0.943*** (0.005)	1.737*** (0.000)
Father's educational level	0.224 (0.342)	0.544** (0.041)
Household's shocks	-0.049 (0.897)	-0.294 (0.498)
Urban community	0.767 (0.421)	0.574 (0.576)
Local unemployment rate	-1.781 (0.383)	-0.405 (0.859)
Local daily wage	0.007 (0.276)	0.007 (0.275)
Constant	-2.113 (0.101)	-5.739*** (0.000)

Note: p-value in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Source: Authors' calculations based on the Young Lives data

Figure 3- 5: Distribution of estimated latent cognitive and non-cognitive skill by education



5.3. Relative importance of skills in labour outcomes

Table 3- 6 provides the estimation results of the equations on labour outcomes. The results presented in the second to the fourth column are the models excluding educational level, and the last three columns are the models with educational level for reference. The second and fifth column provide the results on the probability to be employed or not. The remaining columns give the results on the likelihood of choices on types of employment, which are not employed, wage worker, and self-employed.

Taking into account only skills in the labour outcomes, both cognitive and non-cognitive skills are significant in predicting the employment decisions, with the greater effect size of non-cognitive skill over cognitive skill. greater effect size of non-cognitive skill over cognitive skill.

Estimation results (second to the fourth column) reveal that young people participating in the labour market have lower cognitive and non-cognitive skill levels than the not working group. The probability of not working is higher for the more non-cognitively able individuals than the people with higher cognitive skill (Figure 3- 6). The negative coefficients of skills imply higher skill level of the young people who were students and not working. Of the employment, the more non-cognitively able people tend to be wage workers rather than self-employed. Conversely, the more cognitively able often work as self-employed rather than wage workers (Figure 3- 6).

Table 3- 6: Model of cognitive and non-cognitive skills: labour market participation and employment choice

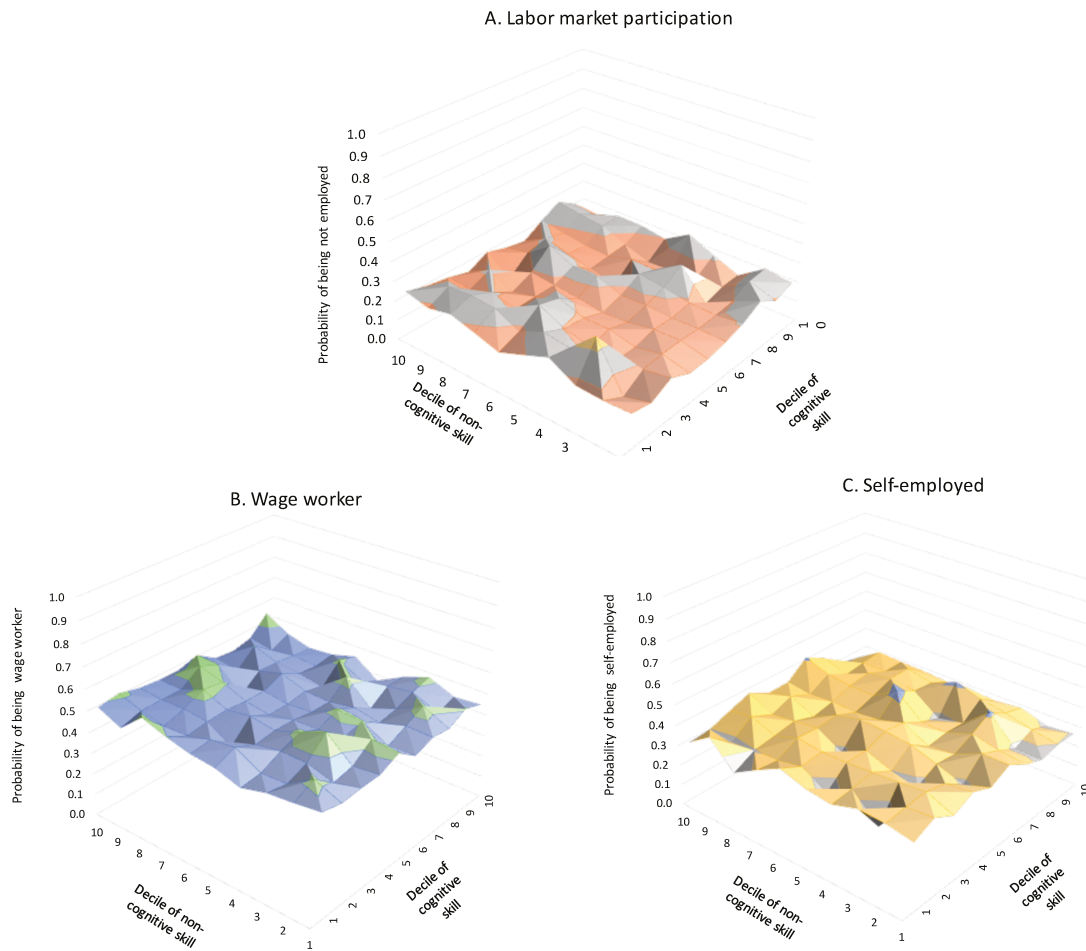
	Employed	Wage worker	Self-employed	Employed	Wage worker	Self-employed
Cognitive skill	-0.474** (0.021)	-0.601*** (0.008)	-0.364* (0.076)	-0.043 (0.844)	0.013 (0.958)	-0.134 (0.554)
Non-cognitive skill	-0.689** (0.034)	-0.578* (0.090)	-0.766** (0.042)	-0.671** (0.041)	-0.526 (0.145)	-0.797** (0.039)
Male	0.488** (0.014)	0.576*** (0.007)	0.408* (0.070)	0.354* (0.082)	0.409* (0.067)	0.334 (0.152)
Lower secondary level				-0.098 (0.839)	-0.277 (0.559)	0.234 (0.644)
Upper secondary level				-1.364*** (0.001)	-1.807*** (0.000)	-0.730 (0.110)
Local unemployment rate	-6.578*** (0.000)	-6.148*** (0.000)	-7.216*** (0.000)	-7.016*** (0.000)	-6.614*** (0.000)	-7.594*** (0.000)
Urban community	-0.062 (0.790)	0.272 (0.260)	-0.651** (0.022)	0.014 (0.954)	0.435* (0.088)	-0.600** (0.041)
Constant	2.642*** (0.000)	1.857*** (0.000)	1.966*** (0.000)	3.744*** (0.000)	3.246*** (0.000)	2.561*** (0.000)

Note: p-value in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Source: Authors' calculations based on the Young Lives data

The significance of the effect of non-cognitive skill on employment decisions is robust regardless of whether the models include or exclude educational variable. The effect sizes of non-cognitive skills across models remain nearly unchanged for the all probabilities of employment choices. Contrarily, the effect of cognitive skill becomes insignificant in the models that additionally take into account educational levels as explanatory variables. In these models, it appears that young people with at least upper secondary education are not likely to enter the labour market. This tendency agrees with the adverse effect of cognitive skills in the model excluding educational level. The analysis of skills and educational choices above reveals that more educated people often have higher cognitive skill. Meanwhile, their higher cognitive skill reduces their probability to take the works. So the significant effect of the cognitive skill on employment choices holds true in these models, but it is unable to define its effect size. The estimates of skill effects suggest the higher predictive power of non-cognitive skill over cognitive skill for employment choices.

Figure 3- 6: Skills and employment choices



In addition to skills, gender and community's characteristic also determine employment choices of young people. Young men are more likely to work, and to be wage workers than young women. Wage worker is more common in the urban areas than in the countryside while self-employed is more likely in the landscape than in the urban areas. In a region where a few young people join the labour market, the probability that a young person works is smaller. This situation suggests either the influential role of the peer in the region on young people's choices on labour market participation. Or the area with a better economic condition or with a high share of better-off families would have lower labour supply and higher demand for pursuing a higher level of education of young people.

The estimation results of earnings equation show that only skills and education determine the earnings capacity of young people. Similar the equations on employment choice, both cognitive and non-cognitive significantly affect earnings. Higher skill levels raise earnings, and the effect size of the non-cognitive skill is greater than the cognitive. The inclusion of the variable on education in the earnings models has no impact on estimated effect of non-cognitive skill but does make the cognitive skill insignificant. Instead, schooling has a significant impact on earnings, with the effect size greater than that of the cognitive skill alone. As education highly correlates with cognitive skill, the estimated effect of educational level in this model includes the effect of cognitive skill. Cognitive thus affects

earnings, but at an undefined effect size. Non-cognitive skill has robust and higher predictive power for earnings. Figure 3- 7 illustrates the level of earnings in response to the increase in skill levels.

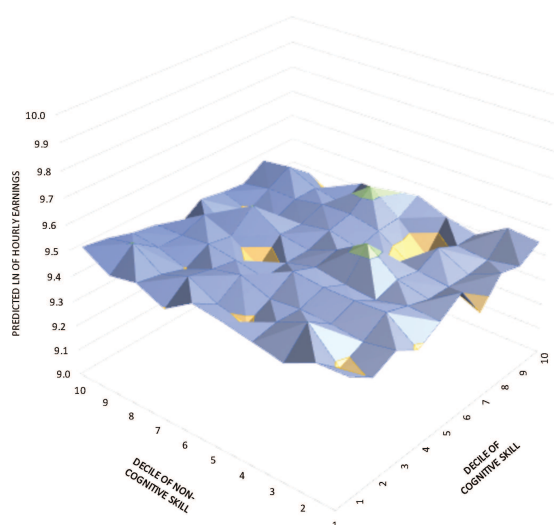
Table 3- 7: Model of cognitive and non-cognitive skills: Log of hourly earnings, whole sample

	Ln hourly earnings	Ln hourly earnings
Cognitive skill	0.206** (0.044)	0.091 (0.249)
Non-cognitive skill	0.290** (0.039)	0.239* (0.096)
Lower secondary level		0.339** (0.018)
Upper secondary level		0.407*** (0.003)
Male	0.090 (0.359)	0.129 (0.199)
Local unemployment rate	0.365 (0.506)	0.478 (0.395)
Urban community (age 19)	0.046 (0.683)	-0.015 (0.900)
Constant	9.310*** (0.000)	8.976*** (0.000)

Note: p-value in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Source: Authors' calculations based on the Young Lives data

Figure 3- 7: Skills and earnings



In sum, both cognitive and non-cognitive skill are significant in determining the labour outcomes of young people. The effect size of non-cognitive is more robust and significant than the cognitive skill. The young people with lower skill levels are more likely to participate in the labour market. Of the employment, the group with higher non-cognitive skill more often to be wage worker, while the more cognitively able group tend to be self-employed.

6. Conclusions and discussion

6.1. Findings

This paper uses the longitudinal data from the Young Lives in Vietnam for the Older Cohort to examine the relative importance of cognitive and non-cognitive skills in labour earnings. The research questions are tackled employing structural equation models, in which latent cognitive and non-cognitive skills and individual, family and community characteristics are allowed to determine the choices on education and employment as well as earnings. Equations on educational and employment choices (probability to be not working/wage worker/self-employed) are estimated using multinomial logit, and labour market participation choice (probability to be employed or not) using logit model, while the linear regression estimates earnings. The paper first investigates the relationship between skills and education, skills and labour outcomes for the whole sample with and without the inclusion of variable on education in the outcome equations. Then the paper tests whether cognitive or non-cognitive skill has higher explanatory power in predicting labour earnings for the young people with the upper secondary level of education.

The estimated latent cognitive skill is more correlated with language skill, while the estimated non-cognitive skill most strongly relates to self-esteem, then peer relations and relates to parent relations the least. Young people with higher cognitive skills tend to obtain a higher level of education. Both cognitive and non-cognitive skills have significant effects on all labour outcomes, with a greater effect size of non-cognitive than cognitive skill. Young people with lower skills are more likely to participate in the labour market. Among the low-skilled individuals, the more cognitively able have, the higher probability to be employed and to work as self-employed, while the more non-cognitive able have a greater chance of being wage workers. Non-cognitive skill raises earnings capacity to a greater extent than the cognitive one.

6.2. Discussion on policy implications

According to the skill formation proposed by Cunha & Heckman (2007, 2008), skills are produced from a combination of individual's abilities, family investment and environment. The environment is a full environment covering home, school, community that affects the skill formation.

The analysis of Heckman et al. (2006) provides the scarce evidence on the causal effect of schooling on non-cognitive skill. Specifically, the locus of control is primarily affected by the high school attendance, while years of education at high school and college affect self-esteem.

International evidence from interventions proves that the non-cognitive skill can be improved in different ages, and by programs with various targets (Almlund et al., 2011). Typical examples are the Perry Program (for poverty reduction in the US) or Jamaica's Supplementation Study (for improving

health) in the early childhood, the Project STAR (for the quality of education in the US) for the middle childhood, and Canadian Self-sufficiency Project (for increasing employability) in the adolescence or early adulthood. Not only by the program focusing on the enhancement of the non-cognitive skill, but the intervention can also be done by the programs that primarily focus on improving the cognitive skill (Perry and STAR) or health status like Supplementation Study. Perry Program aims to enrich the lives of 3-4-year-old low-income black children with IQs less than 85 at age 3. The children attended 2.5-hour center-based preschool five days a week for two years. They were taught how to plan, execute and review a task, and how to work with others when a problem arose. Home visits were conducted to promote parent relations. The program only raised IQ scores until the age 10. It also improved the non-cognitive skills, the locus of control and the self-esteem, which then benefit life outcomes. The Supplementation Study aimed to improve the health of the 1-2-year-old stunted children by providing them two years, either formula milk (nutritional program) or stimulation program (encourage the mother to play with kids) or both. Both interventions boosted cognitive skill, but only stimulation program improves non-cognitive skill in the long term, educational achievement and earnings (by 25 percent at age 22). Project STAR that randomly assigned, from the kindergarten to the third grade, the teachers and children to classrooms with different sizes. Similar to Perry, the positive effect of STAR on achievement test score till the eighth grade. Non-cognitive skills were also improved during the fourth to eighth grade. But the benefits from STAR lasted for a long time later in life, with higher earnings in early adulthood for children sitting in a higher quality kindergarten classroom. Self-sufficiency Project provides single parent earnings subsidies if they found the full-time job and agreed to leave welfare. After 36 months, the subsidy beneficiaries were more likely to have a higher locus of control, higher employment, earnings and less use of welfare. The cost for Perry is USD 11,300 (2007 USD) for 123 participants; STAR: USD 9,355 for 11,571 participants (Cunningham, Acosta, & Muller, 2016).

Numerous papers on the role of early nutrition or health in forming non-cognitive skills for Vietnam, with mixed results, can be referred to in the brainstorming for designing interventions to improve the skill. Georgiadis et al., (2016) verify the positive association between growth in early childhood and later cognitive skills. Dercon & Sanchéz (2011) found that for Vietnam, childhood investment positively predicts child's later self-esteem and educational aspiration and the link is established through a cognitive channel. Additionally, mother's non-cognitive skills prove to be the most important among the predictors of child's non-cognitive skills. The link between early nutrition and skills was then quantified in a similar study by Dercon & Sánchez (2013), with one standard deviation increase in height-for-age tends to increase self-efficacy, self-esteem and aspirations by 10.4%, 6.4% and 5.1%, respectively. Meanwhile, Sanchez (2013) and Le & Tran (2015) show no evidence of the association between height-for-age or growth and non-cognitive skills of adolescents in Vietnam. While these studies as mentioned earlier address the relationship rather than the effect of height-for-age on non-cognitive skills, they suggest the predictive role of nutritional status on later non-cognitive development.

There are also evidences on how other factors, such as family's well-being and child's use of time, feed into forming non-cognitive skills. Dercon & Krishnan (2009) shows the evidence on the link between material well-being and non-cognitive skills, measured by educational aspiration, self-efficacy, self-esteem and inclusion, for the cohort of 11-12 years old in four countries inclusive of Vietnam. Borga (2015) reveals that child's involvement in working is associated with lower level of

self-esteem, self-efficacy of both young children and adolescents in four Young Lives countries inclusive of Vietnam.

The analysis results in this paper reveal the more significant effect non-cognitive over cognitive skill in predicting labour earnings of young people. So, it is also important to improve the non-cognitive skill, especially the self-esteem, for beneficial labour outcomes. The evidence mentioned above from interventions suggests considering two options in developing non-cognitive skills. The first one involves the early investment in child health. The second could be the by-product of the poverty reduction or subsidy programs that target the disadvantaged group such as the low income or less educated parents.

References

- Acosta, P. (2015). Beyond Qualifications: Returns to Cognitive and Socio-Emotional Skills in Colombia. IZA Discussion Paper No. 9403.
- Acosta, P., Muller, N., & Sarzosa, M. (2015). Beyond Qualifications: Returns to Cognitive and Socio-Emotional Skills in Colombia. IZA Discussion Paper No. 9403.
- Almlund, M., Duckworth, A. L., Heckman, J., & Kautz, T. (2011). Personality Psychology and Economics. In *Handbook of the Economics of Education* (Vol. 4, pp. 1–181). Elsevier. Retrieved from <http://linkinghub.elsevier.com/retrieve/pii/B9780444534446000018>
- Bandura, A. (1993). Perceived Self-Efficacy in Cognitive Development and Functioning. *Educational Psychologist*, 28(2), 117–148.
- Bandura, A. (2006). Guide for constructing self-efficacy scales. In *Self-efficacy beliefs of adolescents* (F. Pajares & T. Urdan, Vol. 5, pp. 307–337). Information Age Publishing. Retrieved from <https://www.uky.edu/~eushe2/Bandura/BanduraGuide2006.pdf>
- Blascovich, J., & Tomaka, J. (1991). Measures of Self-Esteem. In *Measure of Personality and Social Psychological Attitudes* (John P. Robinson, Phillip R. Shaver, Lawrence S. Wrightsman, Vol. 1 of Measures of Social Psychological Attitude). Gulf Professional Publishing.
- Borga, L. G. (2015). *Children's own time use and its effect on skill formation*. Prague: Center for Economic Research and Graduate Education.
- Brunello, G., & Schlotter, M. (2011). Non-Cognitive Skills and Personality Traits: Labour Market Relevance and Their Development in Education & Training Systems. *IZA Discussion Paper No. 5743*.
- Carneiro, P., Hansen, K., & Heckman, J. J. (2003). Estimating Distributions of Treatment Effects with an Application to the Returns to Schooling and Measurement of the Effects of Uncertainty on College Choice. *International Economic Review*, 44(2).
- Cunha, F., & Heckman, J. (2007). The Technology of Skill Formation. *American Economic Review*, 97(2), 31–47. <https://doi.org/10.1257/aer.97.2.31>
- Cunha, F., & Heckman, J. (2009). *The Economics and Psychology of Inequality and Human Development* (No. w14695). Cambridge, MA: National Bureau of Economic Research. Retrieved from <http://www.nber.org/papers/w14695.pdf>
- Cunha, F., & Heckman, J. J. (2008). Formulating, Identifying and Estimating the Technology of Cognitive and Noncognitive Skill Formation. *Journal of Human Resources*, 43(4), 738–782. <https://doi.org/10.3368/jhr.43.4.738>
- Cunha, F., Heckman, J., & Schennach, S. (2010). *Estimating the Technology of Cognitive and Noncognitive Skill Formation* (No. w15664). Cambridge, MA: National Bureau of Economic Research. Retrieved from <http://www.nber.org/papers/w15664.pdf>
- Cunningham, W., Acosta, P., & Muller, N. (2016). *Minds and Behaviors at Work: Boosting Socioemotional Skills for Latin America's Workforce*. The World Bank. Retrieved from <http://elibrary.worldbank.org/doi/book/10.1596/978-1-4648-0884-5>
- Cunningham, W., Torrado, M. P., & Sarzosa, M. (2016). *Cognitive and Non-Cognitive Skills for the Peruvian Labor Market: Addressing Measurement Error through Latent Skills Estimations*. The World Bank. Retrieved from <http://elibrary.worldbank.org/doi/book/10.1596/1813-9450-7550>

- Dercon, S., & Krishnan, P. (2009a). Poverty and the Psychosocial Competencies of Children: Evidence from the Young Lives Sample in Four Developing Countries. *Children, Youth and Environments*, 19(2), 138–163.
- Dercon, S., & Krishnan, P. (2009b). Poverty and the Psychosocial Competencies of Children: Evidence from the Young Lives Sample in Four Developing Countries." *Children, Youth and Environments*, 19(2), 138–163.
- Dercon, S., & Sánchez, A. (2011). Long-Term Implications of Under-Nutrition on Psychosocial Competencies: Evidence from Four Developing Countries. Young Lives Working Paper 72. Retrieved from http://www.younglives.org.uk/sites/www.younglives.org.uk/files/YL-WP72_Dercon%26Sanchez_Long-term-implications-of-under-nutrition_oct2011.pdf
- Dercon, S., & Sánchez, A. (2013). Height in mid childhood and psychosocial competencies in late childhood: Evidence from four developing countries. *Economics & Human Biology*, 11(4), 426–432. <https://doi.org/10.1016/j.ehb.2013.04.001>
- Dercon, S., & Singh, A. (2013). From Nutrition to Aspirations and Self-Efficacy: Gender Bias over Time among Children in Four Countries. *World Development*, 45, 31–50. <https://doi.org/10.1016/j.worlddev.2012.12.001>
- Díaz, J. J., Arias, O., & Tudela, D. V. (2013). Does Perseverance Pay as Much as Being Smart? The Returns to Cognitive and Non-cognitive Skills in urban Peru. World Bank Working Paper. Retrieved from <http://econweb.umd.edu/~urzua/DiazAriasTudela.pdf>
- Drago, F. (2011). Self-esteem and earnings. *Journal of Economic Psychology*, 32(3), 480–488. <https://doi.org/10.1016/j.joep.2011.03.015>
- Gallup, J. L. (2002, September). The Wage Labor Market and Inequality in Vietnam in the 1990s. World Bank Policy Research Working Paper #2896. Retrieved from http://www-wds.worldbank.org/external/default/WDSContentServer/IW3P/IB/2002/11/01/000094946_02101805105199/Rendered/PDF/multi0page.pdf
- Georgiadis, A., Benny, L., Crookston, B. T., Duc, L. T., Hermida, P., Mani, S., ... Behrman, J. R. (2016). Growth trajectories from conception through middle childhood and cognitive achievement at age 8 years: Evidence from four low- and middle-income countries. *SSM - Population Health*, 2, 43–54. <https://doi.org/10.1016/j.ssmph.2016.01.003>
- Glewwe, P., Krutikova, S., & Rolleston, C. (2013). Decomposing Gaps in Student Learning in Developing Countries: Evidence from Vietnamese and Peruvian Panel Data. Retrieved from <http://www.ssc.wisc.edu/cde/demsem/YLskldec.pdf>
- Glewwe, P., Qihui, C., & Katare, B. (2012). What Determines Learning Among Kinh and Ethnic Minority Students in Vietnam? An Analysis of the Round 2 Young Lives Data. Young Lives Working Paper 80. Retrieved from <http://younglives.qeh.ox.ac.uk/publications/WP/determines-learning-kinh-ethnic-minority-students-vietnam/wp80-what-determines-learning-among-kinh-and-ethnic-minority-students-in-vietnam>
- Gutman, L. M., & Schoon, I. (2013). The impact of non-cognitive skills on outcomes for young people: Literature review. Institute of Education, University of London.
- Hanushek, E. A. (2014). The Economic Value of Education and Cognitive Skills. In *Handbook of Education Policy Research*. Routledge. Retrieved from <http://www.routledgehandbooks.com/doi/10.4324/9780203880968.ch3>
- Hanushek, E. A., & Woessmann, L. (2008). The Role of Cognitive Skills in Economic Development. *Journal of Economic Literature*, 46(3), 607–668. <https://doi.org/10.1257/jel.46.3.607>

- Heckman, J. J., Humphries, J. E., Urzua, S., & Veramendi, G. (2011). *The effects of educational choices on labor market, health, and social outcomes*. University of Chicago, Department of Economics.
- Heckman, J. J., Lochner, L. J., & Todd, P. E. (2006). Chapter 7 Earnings Functions, Rates of Return and Treatment Effects: The Mincer Equation and Beyond. In *Handbook of the Economics of Education* (Vol. 1, pp. 307–458). Elsevier. Retrieved from <http://linkinghub.elsevier.com/retrieve/pii/S1574069206010075>
- Heckman, J. J., Stixrud, J., & Urzua, S. (2006). The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior. *Journal of Labor Economics*, 24(3), 411–482. <https://doi.org/10.1086/504455>
- Imbert, C. (2013). Decomposing the Labor Market Earnings Inequality: The Public and Private Sectors in Vietnam, 1993-2006. *The World Bank Economic Review*, 27(1), 55–79. <https://doi.org/10.1093/wber/lhs009>
- Krishnan, P., & Krutikova, S. (2013). Non-cognitive skill formation in poor neighbourhoods of urban India. *Labour Economics*, 24, 68–85. <https://doi.org/10.1016/j.labeco.2013.06.004>
- Le, T. D., & Tran, N. T. M. T. (2015). Growth in Middle Childhood and Early Adolescence, and its Association with Cognitive and Non-cognitive Skills at the Age of 15 Years: Evidence from Vietnam. Young Lives Working Paper 130.
- Liu, A. Y. C. (2004). Gender wage gap in Vietnam: 1993 to 1998. *Journal of Comparative Economics*, 32(3), 586–596. <https://doi.org/10.1016/j.jce.2004.04.004>
- Liu, A. Y. C. (2005). Impact of Economic Reform on Wage Structure in Vietnam. *East Asian Economic Perspective*, 16(2), 46–57.
- Liu, A. Y. C. (2008). Changes in urban inequality in Vietnam: 1992–1998. *Economic Systems*, 32(4), 410–425. <https://doi.org/10.1016/j.ecosys.2007.10.002>
- Moock, P. R., Patrinos, H. A., & Venkataraman, M. (2003). Education and earnings in a transition economy: the case of Vietnam. *Economics of Education Review*, 22(5), 503–510. [https://doi.org/10.1016/S0272-7757\(03\)00039-6](https://doi.org/10.1016/S0272-7757(03)00039-6)
- MPI-GSO, M. of P. and I.-G. S. O. (2016). *Năng suất lao động của Việt Nam: Thực trạng và giải pháp [Productivity in Vietnam: situation and solution]*.
- Murnane, R. J., Willett, J. B., Braatz, M. J., & Duhaldeborde, Y. (2001). Do different dimensions of male high school students' skills predict labor market success a decade later? Evidence from the NLSY. *Economics of Education Review*, 20, 311–320.
- Nguyen, L. D., & Thi Tran, H. (2016). Determinants to Female Labor Income in the Vietnam's Informal Sector. *Mediterranean Journal of Social Sciences*. <https://doi.org/10.5901/mjss.2016.v7n3s1p169>
- Nordman, C. J., Sarr, L. R., & Sharma, S. (2015). Cognitive, Non-Cognitive Skills and Gender Wage Gaps: Evidence from Linked Employer-Employee Data in Bangladesh. IZA Discussion Paper No. 9132.
- OECD. (2015). *Skills for Social Progress*. OECD Publishing. Retrieved from http://www.oecd-ilibrary.org/education/skills-for-social-progress_9789264226159-en
- Oostendorp, R. H., & Doan, Q. H. (2013). Have the returns to education really increased in Vietnam? Wage versus employment effect. *Journal of Comparative Economics*, 41(3), 923–938. <https://doi.org/10.1016/j.jce.2012.12.002>

- Pham, T.-H., & Reilly, B. (2007). The gender pay gap in Vietnam, 1993–2002: A quantile regression approach. *Journal of Asian Economics*, 18(5), 775–808. <https://doi.org/10.1016/j.asieco.2007.04.001>
- Pham, T.-H., & Reilly, B. (2009). Ethnic wage inequality in Vietnam. *International Journal of Manpower*, 30(3), 192–219. <https://doi.org/10.1108/01437720910956727>
- Phan, D., & Coxhead, I. (2013). Long-run costs of piecemeal reform: Wage inequality and returns to education in Vietnam. *Journal of Comparative Economics*, 41(4), 1106–1122. <https://doi.org/10.1016/j.jce.2013.04.001>
- Rolleston, C., & James, Z. (2011). The Role of Schooling in Skill Development- evidence from Young Lives in Ethiopia, India, Peru and Vietnam. Background paper for the Education for All Global Monitoring Report 2012.
- Rolleston, C., James, Z., Tran, N. T. M. T., & Pasquier-Doumer, L. (2013). Making Progress: Report of the Young Lives School Survey in Vietnam. Young Lives Working Paper 100. Retrieved from http://www.younglives.org.uk/publications/WP/making-progress-report-of-the-young-lives-school-survey-in-vietnam/yl-wp100_rolleston
- Rosen, J. A., & RTI International (Eds.). (2010). *Noncognitive skills in the classroom: new perspectives on educational research*. Research Triangle Park, NC: RTI Press.
- Rosenberg, M. (1965). Society and the adolescent self-image. *Princeton University Press*.
- Roseth, V. V., Valerio, A., & Gutiérrez, M. (2016). *Education, skills, and labor market outcomes: results from large-scale adult skills surveys in urban areas in 12 countries* (Main report). Washington, D.C.: World Bank Group. Retrieved from <http://documents.worldbank.org/curated/en/397871467995623739/Main-report>
- Rotter, J. B. (1966). Generalized expectancies for internal versus external control of reinforcement. *Psychological Monographs: General and Applied*, 80(1).
- Sahn, D. E., & Villa, K. M. (2015). The Role of Personality, Cognition and Shocks in Determining Age of Entry into Labor Market, Sector of Employment, and within Sector Earnings (Vol. No 205673). Presented at the 2015 AAEA & WAEA Joint Annual Meeting, July 26-28, San Francisco, California, Agricultural and Applied Economics Association. Retrieved from <http://purl.umn.edu/205673>
- Sakellariou, C. (2009). Endogeneity, computers, language skills and wages among university graduates in Vietnam. *Applied Economics*, 41(5), 653–663. <https://doi.org/10.1080/00036840601007377>
- Sakellariou, C., & Fang, Z. (2014). The Vietnam reforms, change in wage inequality and the role of the minimum wage. *Economics of Transition*, 22(2), 313–340. <https://doi.org/10.1111/ecot.12037>
- Sanchez, A. (2013). The Structural Relationship between Nutrition, Cognitive and Non-cognitive Skills Evidence from four developing countries. Young Lives Working Paper 111.
- Schwarzer, R., & Jerusalem, M. (1995). Optimistic Self-Beliefs as a Resource Factor in Coping with Stress. In S. E. Hobfoll & M. W. Vries (Eds.), *Extreme Stress and Communities: Impact and Intervention* (pp. 159–177). Dordrecht: Springer Netherlands. Retrieved from http://link.springer.com/10.1007/978-94-015-8486-9_7
- Singh, A. (2014). Emergence and evolution of learning gaps across countries: Linked panel evidence from Ethiopia, India, Peru and Vietnam. CSAE Working Paper WPS/2014-28. Retrieved from <http://www.csae.ox.ac.uk/workingpapers/pdfs/csae-wps-2014-28.pdf>

World Bank. (2014). *Skilling up Vietnam: Preparing the workforce for a modern market economy* (No. 82940 V2). Vietnam: World Bank. Retrieved from http://www-wds.worldbank.org/external/default/WDSContentServer/WDSP/IB/2013/11/26/000461832_20131126115640/Rendered/PDF/829400AR0P13040Box0379879B00PUBLIC0.pdf

Appendices

A. Summary on the evidences on skills and labour outcomes

Paper	Country	Data	Direct/ indirect effect of skills on labour outcomes	Measurement of cognitive skills	Measurement of non- cognitive skills	Measure- ment of labour outcomes	Causal identifi- cation	Treat- ment for reverse causality	Treat- ment for reverse causality	Treat- ment for endoge- neity	Results
Murane et al (2001)	developed (US)	longitudinal (NLSY79)	direct	2 measures: academic score (composite score on 4 subtests of ASVAB: Arithmetic Reasoning, Word Knowledge, Paragraph Comprehension, Mathematics Knowledge), speed score (subtests of the ASVAB: Numerical Operations and Coding Speed)	1 measure: Self- esteem (composite scores on 7 out of Rosenberg 10-item Self- Esteem Scale)	wages at age 27 and 28	no	yes	no	no	1 SD increase in academic skills and self-esteem is associated with nearly 10 and 4 % increase in wage respectively
Drago (2011)	developed (US)	longitudinal (NLSY79)	direct		1 measure: Self- esteem (aggregate score based on Rosenberg 10- item Self- Esteem Scale)	wages	no	yes	no	no	1 SD in self-esteem in 1987 is associated with from 7.9% increase in wage

Paper	Country	Data	Direct/ indirect effect of skills on labour outcomes	Measurement of cognitive skills	Measurement of non- cognitive skills	Measure- ment of labour outcomes	Causal identifi- cation	Treat- ment for reverse causality	Treat- ment for reverse causality	Treat- ment for endoge- neity	Results
Heckman et al (2006)	developed (US)	longitudinal (NLSY79)	direct	1 latent measures (based on 5 test scores: arithmetic reasoning, word knowledge, paragraph comprehension, mathematical knowledge, coding speed)	1 latent measure (based on 2 scores on Rotter Internal- External Locus of Control Scale and Rosenberg Self-Esteem Scale)	labour force participati on; white/blu e collar, years of (working) experien ce, wages	yes	yes	yes	yes	<ul style="list-style-type: none"> - Cognitive skills have about the same effect on wages as non-cognitive skills; - The effects of non-cognitive skills on employment is greater than cognitive skills; cognitive skills have slightly larger effects on the choice of white versus blue collar occupations than non-cognitive skills; - Effects of skills vary across gender and schooling.
Heckman et al (2011)	developed (US)	longitudinal (NLSY79)	indirect (through education)	1 latent measure (based on 5 test scores: arithmetic reasoning, word knowledge, paragraph comprehension, numerical operations, coding speed)	1 latent measure (based on adolescent behavioural outcomes: risky; illegal behaviours; high school club member)	labour force participati on; white/blu e collar, wages	yes	yes	yes	yes	Significant gains in labour market outcomes from graduating high school and going to college, and the gains are larger for high ability group

Paper	Country	Data	Direct/ indirect effect of skills on labour outcomes	Measurement of cognitive skills	Measurement of non- cognitive skills	Measure- ment of labour outcomes	Causal identifi- cation	Treat- ment for reverse causality	Treat- ment for reverse causality	Treat- ment for endoge- neity	Results
Díaz et al. (2013)	developing (Peru)	cross- section	direct	5 measures: Peabody's Picture Vocabulary Test, verbal fluency, working memory, math problem- solving; and aggregate measure (obtained from these 4 tests)		earnings	no	no	no	yes	- 1 SD change in both overall cognitive skill and perseverance facet of Grit has the equal effect size of 9% on earnings, conditional on schooling. - Lower returns to earnings are observed for agreeableness (8%) emotional stability (5%)
Krishnan and Krutikova (2013)	developing (India)	cross- section	direct		2 measures: self-esteem, self-efficacy	wage	no	no	no	no	Self-efficacy and self-esteem are positively correlated with wage
Nordman et al. (2015)	developing (Bangladesh)	cross section	direct	literacy, numeracy	7 measures: Big Five personality traits; Hostile bias, decision making	wage	no	no	no	no	- Reading and numeracy: 1 SD increase in reading and numeracy is associated with around 5% and 4% increase in mean wage; in high correlation with wages at lower percentiles - Personality traits: insignificant to mean wage;

Paper	Country	Data	Direct/ indirect effect of skills on labour outcomes	Measurement of cognitive skills	Measurement of non- cognitive skills	Measure- ment of labour outcomes	Causal identifi- cation	Treat- ment for reverse causality	Treat- ment for reverse causality	Treat- ment for endoge- neity	Results
											<p>nuanced effect of non-cognitive skills on mean wage across conditional wage distribution: Agreeableness is positively associated with wage at the 10th, 25th and 50th percentiles; Emotional Stability positively associated with wage at 1st decile; Openness negatively associated with wages at 10th percentiles; Conscientiousness negatively associated with wages at 10th percentile, positively associated with wages at 90th percentile.</p>
Sahn & Villa (2015)	developing (Madagascar)	cross-section (longitudinal for community characteristics and local school)	direct	1 measure (standardized aggregate score of two math and two French tests)	5 measures (Openness to Experience, Conscientiousness, Extraversion, Agreeableness, & Neuroticism)	labour market participation; employment; employment sector; earnings	no	no	no	yes	<p>- Earnings: negative effect of cognitive skill; positive effect of Conscientiousness and negative effect of Neuroticism on male (female) earnings in informal (formal) sector; mixed effects of Extraversion and Openness</p>

Paper	Country	Data	Direct/ indirect effect of skills on labour outcomes	Measurement of cognitive skills	Measurement of non- cognitive skills	Measure- ment of labour outcomes	Causal identifi- cation	Treat- ment for reverse causality	Treat- ment for reverse causality	Treat- ment for endoge- neity	Results
											<ul style="list-style-type: none"> - Labour market participation: negative (positive) effect of Openness (Neuroticism) on the hazard of entering the labour market after age 10 (but positive effect after age 14) - Male employment: 1 SD increase in: Neuroticism associated with 3.2% increase in probability of being unemployed; Openness and Extraversion - 5.7% and 6% in likelihood of being employed in formal sector; Extraversion- 5.8% in likelihood of being employed in informal sector
Acosta et al. (2015)	developing (Columbia)	cross-section		1 latent measure based on 3 test scores: language; use and length of reading on and off work; reading proficiency	1 latent measure based on 3 aggregate scores: (1) Extraversion and Openness; (2) Emotional Stability, Hostile bias and Concienciousness;	earnings, labour force participation, occupational choices	no	no	yes	yes	<ul style="list-style-type: none"> - Cognitive skill is positively associated with earnings (13.4%); the probability of being formal worker (27.6%); high skilled worker (25.2%); - Non-cognitive skill has insignificant effects on labour outcomes

Paper	Country	Data	Direct/ indirect effect of skills on labour outcomes	Measurement of cognitive skills	Measurement of non- cognitive skills	Measure- ment of labour outcomes	Causal identifi- cation	Treat- ment for reverse causality	Treat- ment for reverse causality	Treat- ment for endoge- neity	Results
					(3) Grit and Decision making)						
Cunningham et al. (2016)	developing (Peru)	cross section	direct	1 latent measures for cognitive skill (based on 4 achievement tests Peabody Picture Vocabulary test, verbal fluency, working memory, numeracy/problem- solving)	2 latent measures for non-cognitive skill: Stability (based on Agreeableness, Conscientious- ness, Neuroticism); Plasticity (based on Extraversion, Openness); and Grit (based on perseverance of effort, persistence of goals, consistency of interest)	wage, employment choice	no	no	yes	yes	<ul style="list-style-type: none"> - Cognitive skill is positively associated with wage (13%), the probability of being wage worker (27%), white collar worker (31%) and worker in the formal sector (22%). - Individual in the 10th decile of Plasticity is 35% more likely to be employed; - Individual in the 10th decile of Grit is 45% and 20% more likely to be wage worker and to hold a white-collar job than the 1st decile.

B. Tables

B.1. Factor analysis for cognitive and non-cognitive skill index

Cognitive skill index

Factor analysis (principal component factors; 1 factor retained; 812 obs)

Factor	Eigenvalue	Difference	Proportion	Cumulative
Factor 1	1.61322	1.22645	0.8066	0.8066
Factor 2	0.38678	.	0.1934	1

LR test: independent vs. saturated: $\chi^2(1) = 382.29$ Prob> $\chi^2 = 0.0000$

Factor loading and unique variances

Variable	Factor 1	Uniqueness
Math	0.89812	0.19339
Language	0.89812	0.19339

Estimation sample factor

Variable	Mean	Std. Dev.	Min	Max
Math	-0.0019355	0.5305851	-3.504867	2.099888
Language	0.0042966	0.5719565	-3.433159	2.403103

(Factor analysis correlation matrix)

Non-cognitive skill index

Factor analysis (Principal component factors; 1 factor retained; 854 obs)

Factor	Eigenvalue	Difference	Proportion	Cumulative
Factor 1	2.23732	1.42189	0.5593	0.5593
Factor 2	0.81543	0.25993	0.2039	0.7632
Factor 3	0.5555	0.16375	0.1389	0.9021
Factor 4	0.39175	.	0.0979	1

LR test: independent vs. saturated: $\chi^2(6) = 786.90$ Prob> $\chi^2 = 0.0$

Factor loading and unique variances

Variable	Factor 1	Uniqueness
Self-efficacy	0.7464	0.44288
Self-esteem	0.83566	0.30167
Parent relations	0.60587	0.63292
Peer relations	0.78409	0.3852

Estimation sample factor

Variable	Mean	Std. Dev.	Min	Max
Self-efficacy	-0.0019355	0.5305851	-3.504867	2.099888
Self-esteem	0.0042966	0.5719565	-3.433159	2.403103
Parent relations	-0.001561	0.6597391	-3.841397	1.421762
Peer relations	0.0029838	0.6057027	-3.512921	2.137697

(Factor analysis correlation matrix)

B.2. List of explanatory variables included in outcome equations

Explanatory variables	Dependent variables (model)			
	Skill factors (linear)	Educational choices (multinomial logit)	Choices on employment (logit, multinomial logit)	Earnings (linear)
Cognitive skill	yes	yes	yes	yes
Non-cognitive skill	yes	yes	yes	yes
Male	yes	yes	yes	yes
Number of siblings	yes	yes		
Wealth index	yes	yes		
Mother's educational level	yes	yes		
Father's educational level	yes	yes		
Household's shocks	yes	yes		
Local daily wage	yes	yes		
Local unemployment rate	yes	yes	yes	yes
Urban community	yes	yes	yes	yes

B.3. Tests for normality of estimated latent skills

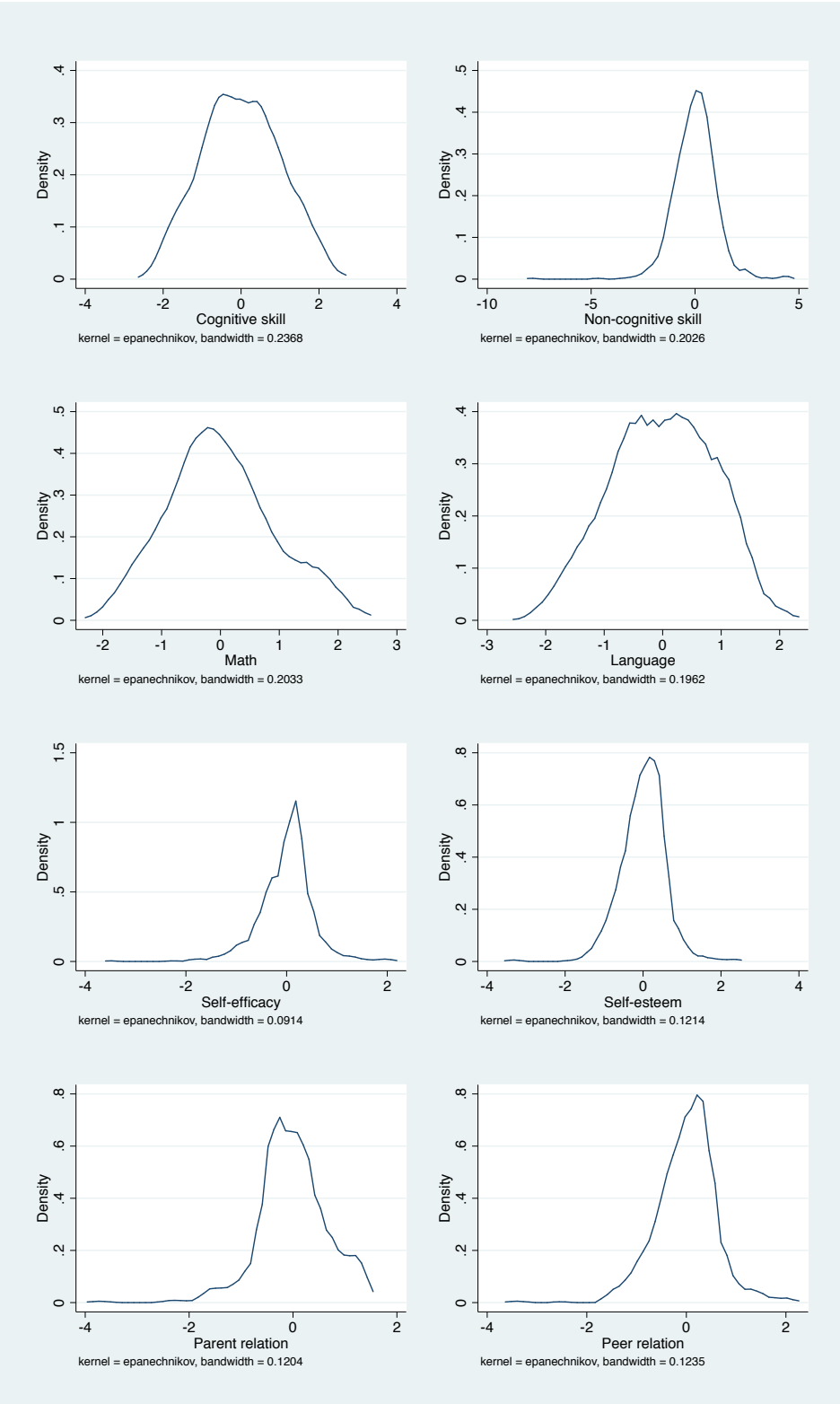
	Shapiro-Wilk W test	Shapiro-Francia test	Skewness/Kurtosis tests
Cognitive skill	0.008	0.023	0.004
Non-cognitive skill	0.000	0.000	0.000

Note: the figures are statistics for the mentioned tests for normality of the estimated latent skills. The null hypothesis is the estimated skills are normally distributed.

Source: Authors' calculations based on the Young Lives data

C. Figures

C.1. Distribution of skills



Résumé de la thèse

"Formation des compétences et transition vers des emplois productifs au Vietnam"

1. Le contexte du Vietnam

Réussite exceptionnelle dans le domaine de l'éducation et de l'acquisition de compétences cognitives

Le Vietnam a obtenu des succès significatifs en matière de développement économique et de réduction de la pauvreté au cours des deux dernières décennies. Entre 1993 et 2008, le taux de croissance annuel moyen du PIB était de 7.2%. Le taux de pauvreté a fortement diminué de 58% à environ 15% au cours de cette période. Le taux de croissance élevé au cours de la période n'a pas été accompagné d'une hausse des inégalités. À côté de la croissance économique, le secteur de l'éducation au Vietnam a connu un développement important. Par rapport aux pays de même niveau de revenu, le nombre d'années de scolarité des adultes de ce pays est le plus élevé. Les taux de scolarisation à tous niveaux ont augmenté de façon spectaculaire pour atteindre des niveaux plus élevés que d'autres pays à faible revenu. En ce qui concerne l'enseignement de base, la scolarisation primaire universelle est presque atteinte avec un taux net de 99% au cours de l'année scolaire 2013-2014. De plus, le taux net de scolarisation dans l'enseignement secondaire a continué à augmenter de 70 à 90% entre 2000 et 2014.

La réussite la plus remarquable de l'éducation est l'importance de l'acquisition d'aptitudes cognitives par rapport aux autres pays de même niveau de revenu ou de revenu beaucoup plus élevé. La compétence cognitive qui désigne la capacité de résoudre les problèmes abstraits est mesurée par les résultats des tests cognitifs tels que le test de QI ou de mathématiques. Un programme international a été mené par l'OCDE, visant à évaluer les résultats des élèves de 15 ans sur plusieurs compétences. Selon ces études, parmi les 65 pays participants en 2012, le Vietnam s'est classé 8ème en sciences, 17ème en mathématiques et 19ème en lecture. En 2014, les résultats correspondants étaient de 8ème, 22ème et 32ème parmi les 72 pays participants. Sur le classement de toutes les matières, le Vietnam se situait au-dessus de la moyenne de l'OCDE. Si la compétence de lecture n'est pas comptée, en 2015, il est placé au-dessus des Etats-Unis, l'Australie, le Royaume-Uni et la France. D'après les résultats des enquêtes STEP66, la main-d'œuvre vietnamienne occupe la première place de l'alphabétisation et la deuxième dans l'intensité d'utilisation des compétences en lecture et en écriture parmi les pays en développement. En termes de compétence cognitive, le Vietnam est un exemple qui mérite d'être multiplié dans d'autres pays.

Mais des écarts de compétences et de faibles compétences non cognitives

Il existe cependant des limites à ce développement. Malgré de faibles changements dans les inégalités, il subsiste des disparités dans les résultats et les opportunités entre les différents segments de la société. L'inégalité des revenus a augmenté dans la moitié supérieure de la distribution et entre les régions rurales et

⁶⁶Programme de Banque mondiale pour mesurer la compétence en matière de capacité d'insertion professionnelle et de productivité - qui consiste en une enquête auprès des ménages et une enquête auprès des employeurs. Le programme a recueilli des informations sur l'offre et la demande de compétences dans les zones urbaines de 12 pays en développement, dont la Bolivie, la Colombie, le Ghana, la République démocratique populaire Lao, l'Ukraine, le Vietnam et la province du Yunnan en Chine et en Arménie, Azerbaïdjan, Géorgie et Kenya. ARY de Macédoine en 2013.

urbaines. Les différences entre les ethnies minoritaires et les Kinh ont augmenté, notamment en termes de revenu dans les zones rurales.

De plus, les disparités géographiques - entre les zones rurales et urbaines / provinces / zones rurales et urbaines au sein d'une province ont augmenté. Les inquiétudes de l'opinion publique concernant la montée des inégalités et des inégalités dans l'accès à l'éducation se font plus virulentes. En 2012, plus de la moitié des habitants dans les zones rurales et 76% des citoyens étaient préoccupés par l'inégalité de niveau de vie. Notamment, le groupe rural est celui qui est le plus préoccupé par la qualité et l'accès aux services de santé, tandis que la qualité et l'accès à l'éducation sont les problèmes des urbains. Au Vietnam, les déficits d'apprentissage sont liés à l'origine sociale et prennent place à un âge précoce, avant l'âge de la scolarité formelle.

Avec les inégalités en matière de bien-être et d'opportunités, il existe des différences dans la participation à l'éducation, le niveau d'éducation et la réussite entre les régions et les niveaux de revenus / de dépenses. L'inégalité du niveau de scolarité est relativement faible avec un coefficient de Gini de 0,25. Mais ce coefficient varie fortement entre les provinces et les ethnies ces importantes entre les coefficients provinciaux et ethniques de Gini (de 0,18 à 0,53) et les années de scolarité (de 4 à 11 ans).

Les inégalités d'accès à l'éducation sont faibles pour le primaire mais sont importantes pour le premier et le deuxième cycle du secondaire. En 2012, le taux net de scolarisation secondaire en premier et en deuxième cycle pour ceux qui avaient les revenus les plus élevés atteignait 94% et 83%. Cependant, les parts pour ceux de quintiles de revenu le plus bas étaient de 71 et 36%. De même, les enfants des groupes ethniques minoritaires avaient en 2014 des taux net de scolarisation dans l'enseignement secondaire inférieur et supérieur de 10 et 30 points de pourcentage inférieurs à ceux de l'ethnie majoritaire. Parmi ceux de faible revenu, les minorités et les résidents ruraux avaient des niveaux d'éducation plus bas que leurs homologues des résidents à revenu élevé, majoritaires et urbains. En ce qui concerne les résultats d'apprentissage, les filles, les enfants dont la famille possède un niveau de richesse et d'éducation supérieur avaient de meilleurs résultats (en mathématiques et en vietnamien) que les garçons et les enfants dont la famille est moins riche.

Une autre source de l'écart de compétence est la tendance du pays en matière de décrochage scolaire. Le Vietnam comptait parmi les dix pays du monde avec plus d'un million d'enfants non scolarisés en 2005. Le nombre d'élèves quittant l'école a augmenté fortement ces dernières années. En 2009, 1,26 million d'enfants de 5 à 14 ans ont abandonné l'école ou n'avaient jamais fréquenté l'école (pré-primaire / primaire / secondaire). Si l'on exclut les jeunes qui quittent l'école au deuxième cycle de l'enseignement secondaire (9e année), le décrochage scolaire au Vietnam persiste à environ 11%. Comme le premier cycle de l'enseignement secondaire est un niveau de scolarité qui fournit aux élèves des compétences de base en lecture et en calcul, les jeunes qui quittent l'école à ce niveau manquent de compétences de longue durée pour poursuivre leurs études et de gagner leur vie.

Les écarts de compétences provoquent un cercle vicieux de pauvreté. Les enfants nés dans certaines circonstances, par exemple les enfants des ethnies minoritaires ou les enfants dont les parents sont les moins éduqués ont une demande plus faible d'éducation. À cet égard, l'inégalité des chances d'acquérir une bonne éducation et des compétences adéquates peut exacerber les inégalités auxquelles sont confrontés les enfants et perpétuer leur pauvreté.

Deuxièmement, une faiblesse importante de l'éducation est son incapacité à fournir aux étudiants des compétences non cognitives adéquates. La compétence non cognitive est définie par des traits de personnalité ou des capacités faiblement corrélées avec des mesures d'intelligence. Contrairement à la performance impressionnante en matière de compétences cognitives, la main-d'œuvre vietnamienne affiche un faible niveau de compétences non cognitives parmi les pays en développement. Comme le montrent les résultats des

enquêtes auprès des ménages STEP, les scores du Vietnam pour l'extraversion, l'ouverture, l'agrément, la prise de décision et les préjugés hostiles se situent dans le tiers inférieur et la conscience et la gravité dans les deuxièmes et troisièmes. La tendance exclut la stabilité émotionnelle qui est la plus élevée pour les Vietnamiens. L'analyse de l'enquête auprès des employeurs de STEP montre également que la plupart des employeurs sont confrontés au manque de compétences requises et à l'indisponibilité de candidats qualifiés lors de la procédure d'embauche.

Le paradoxe

À l'heure actuelle, malgré une croissance de la productivité du travail assez élevée par rapport aux autres pays de la région, la productivité du travail au Vietnam est une des plus faibles de la région Asie-Pacifique. Cela pose la question de ce qui se cache derrière le paradoxe d'une productivité faible mais de très haut niveau de compétence cognitive et de réussite scolaire. Une explication possible de cette question est la faiblesse du niveau de compétence non cognitive. Les compétences de la main-d'œuvre constituent un capital humain qui est, à son tour, une des sources principales de productivité du travail. En raison du développement rapide de la technologie et de l'intégration économique du monde, la demande de travailleurs qualifiés augmente tandis que la demande de travailleurs moins qualifiés diminue à l'échelle mondiale. Le Vietnam n'est pas une exception, en particulier dans le contexte d'une transition d'une économie agricole vers une économie industrialisée moderne en 2020 (Stratégie de développement socioéconomique du Vietnam pour la période 2011-2020). Alors que la main-d'œuvre acquiert des aptitudes cognitives élevées, la demande pour des compétences hautement non cognitives n'a pas été satisfaite. La compétence non cognitive a potentiellement un rendement plus élevé que celle cognitive sur les résultats du marché du travail, ce qui entraîne une faible productivité du pays. Une autre raison de la faible productivité pourrait être dans les niveaux d'inégalité et d'inégalité des chances. De nombreux éléments de preuve impliquent qu'au-delà d'un petit seuil d'inégalité des revenus, un niveau d'inégalité plus élevé affecte négativement la productivité. Dans une perspective plus détaillée, l'augmentation de l'inégalité augmente le pouvoir de décision pour la classe supérieure dans la société. Plus la disparité est élevée, plus la classe supérieure domine dans tous les aspects de la vie. Ils ont la possibilité de manipuler le système (le marché, les médias, le gouvernement, etc.) pour acquérir des choses dans leur intérêt et gagner plus de revenus. L'inégalité accrue et l'inégalité des chances conduisent ainsi à un revenu plus élevé de la classe supérieure d'une manière improductive, en d'autres termes, les inégalités viennent avec une productivité réduite. Étant donné le vieillissement rapide de la population vietnamienne et le ralentissement continu d'économie, la productivité faible du travail est préjudiciable à la croissance économique du pays dans les années à venir ainsi que la perspective à long terme du développement du Vietnam.

De la discussion ci-dessus, les inégalités dans l'éducation, les écarts de compétences et l'insuffisance des compétences non cognitives suscitent l'inquiétude sur les résultats ultérieurs du capital humain, la croissance économique et les inégalités au Vietnam. Les sources d'inégalité en matière d'éducation et les résultats en matière de réussite cognitive et non cognitive ne sont pas encore claires. Les limites du système éducatif que sont la prévalence du tutorat privé, l'insuffisance du salaire officiel d'un enseignant pour couvrir ses dépenses de vie, un système d'éducation axé sur les examens plutôt que d'autres aspects des compétences des élèves la corruption et l'abus de pouvoir peuvent être sources de ces inégalités. Ces limites ne garantissent pas la qualité de l'apprentissage et faussent les résultats d'apprentissage. Chacun de ces obstacles peut, seul ou en combinaison avec d'autres, avoir une incidence négative sur l'acquisition de compétences et les lacunes dans les compétences. Par exemple, la faible rémunération des enseignants, le tutorat privé, la corruption et l'abus de pouvoir pourraient aggraver l'inégalité. Ou le système d'éducation axé sur les examens négligent les compétences non cognitives, soit en améliorant les compétences cognitives pour les enfants qui suivent des cours privés ou ne s'améliore pas; à la place, les résultats des tests sont exagérés en raison de la corruption et de l'abus de pouvoir. Une rémunération insuffisante pour les enseignants encourage le tutorat privé des

enseignants ou la corruption et l'abus de pouvoir des enseignants, affectant ainsi la qualité de la scolarité et l'inégalité.

Pour ces faits, il apparaît nécessaire de réformer de manière significative le secteur de l'éducation, principalement en ce qui concerne la qualité des écoles générales, afin d'assurer l'égalité des chances dans l'éducation des enfants. Il y a la demande des décideurs d'identifier les obstacles à l'acquisition de types et de niveaux de compétences adéquats. Il est plus urgent d'étudier le processus de formation des compétences pour bien comprendre comment les facteurs innés et environnementaux (la famille, l'école et la communauté) interagissent sur l'enfance pour influencer les résultats de l'enfant en matière de compétences et de réussite scolaire.

2. Recherches de compétences et des inégalités éducatives pour Vietnam

Les recherches existantes sur la formation des compétences, les inégalités dans l'éducation et le bien-être au Vietnam font face à trois principales limites. Premièrement, la plupart des études ont porté sur l'établissement de la corrélation entre les antécédents sociaux et les résultats scolaires, ou l'environnement scolaire et les résultats scolaires, mais sans aborder l'interaction entre le milieu social et l'environnement scolaire ou les liens entre ces trois composantes. La raison de cette limite est que les études existantes reposent soit sur des enquêtes auprès des ménages qui fournissent des informations détaillées sur le milieu social mais très peu d'informations sur l'environnement scolaire. Ou les études utilisent des données provenant d'enquêtes scolaires qui mesurent les compétences cognitives et l'environnement scolaire, mais donnent peu d'informations sur le contexte social, car les jeunes enfants ou les enseignants sont les répondants. La deuxième limite est qu'ils ne tiennent pas compte du rôle des compétences non cognitives dans le processus éducatif et des compétences cognitives et non cognitives dans la transmission des inégalités. La troisième limite concerne l'utilisation des notes à des tests comme indicateurs de compétences dans les études mentionnées ci-dessus, qui est associée à une erreur de mesure parce que les résultats des tests révèlent non seulement des compétences mais aussi d'autres facteurs tels que les efforts ou les préférences.

Pour ces limitations, il n'y a pas eu de recherche sur la façon dont la formation des compétences détermine les résultats sur l'éducation et le revenu ou le salaire au Vietnam. Au lieu de cela, aucune preuve du lien entre les compétences et les résultats sur le marché du travail n'existe pour le Vietnam. Ni la preuve sur la façon dont les compétences et les lacunes dans les compétences déterminent l'inégalité dans le revenu ou le salaire.

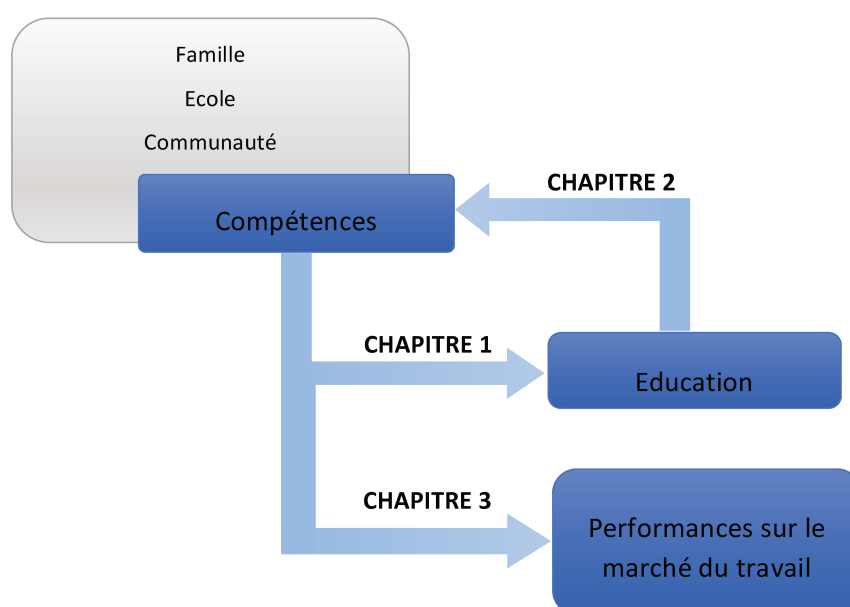
La littérature sur le Vietnam fournit peu d'éléments pour comprendre le paradoxe de compétence cognitive élevée accompagnée d'une faible productivité du travail. Dans ma thèse, je vise à apporter des éclairages sur ce thème en utilisant une base de données originale collectée par mon institution de recherche, le Centre d'analyse et de prévision dans le cadre du projet Young Lives. Le projet Young Lives est une étude à long terme de la pauvreté infantile pour comprendre ses causes et conséquences afin de formuler des politiques publiques. Il est dirigé par une équipe du Département du développement international de l'Université d'Oxford. En association avec d'autres partenaires, les données sont collectées dans quatre pays: l'Ethiopie, l'Inde (Andhra Pradesh), le Pérou et le Vietnam. Il suit sur 15 ans deux cohortes d'enfants, la cohorte jeune et la cohorte plus âgée, née en 2001-02 et 1994-95 respectivement. Dans ce projet, des enquêtes auprès des ménages sont menées tous les quatre ans et sont combinées avec une enquête scolaire menée à l'âge de dix ans. Ces données combinent ainsi des informations sur l'origine sociale des enfants, sur leur expérience scolaire, sur les caractéristiques de l'école et de l'enseignant, ainsi que sur leurs connaissances cognitives et non cognitives, dans une perspective dynamique. Profitant de ces données, ma thèse vise à mieux comprendre la façon dont les compétences interagissent avec les facteurs environnementaux pour déterminer les résultats scolaires et les moyens de subsistance au Vietnam. Premièrement, mon analyse des résultats scolaires et du travail tient compte à la fois des compétences cognitives et non cognitives. Deuxièmement, je cherche à étudier les mécanismes sous-jacents

aux résultats scolaires grâce à la combinaison des enquêtes auprès des ménages et des écoles. En troisième lieu, j'aborde les erreurs de mesure des compétences en combinant des résultats des tests et d'autres facteurs.

3. Objectif, questions de recherche et chapitres

Ma thèse se concentre sur les enfants au Vietnam et identifie les barrières aux opportunités éducatives de ces enfants. Cette recherche devrait permettre de renseigner une réforme de l'éducation visant la qualité de l'école pour tous. Il cherche à examiner les effets des caractéristiques scolaires et du contexte social sur la formation professionnelle et la transition vers des moyens de subsistance productifs en abordant le lien tridimensionnel entre les compétences, l'éducation et les résultats du travail (voir la figure ci-dessous). Dans la thèse, j'aborderai trois sous-questions comme ci-dessous.

Articulation de la thèse



1) Comment la compétence décide le décrochage aux adolescents?

Comme discussion ci-dessus, le Vietnam a obtenu des résultats exceptionnels en l'éducation et l'acquisition de compétences. Cependant, il y a eu quelques problèmes émergents du secteur de l'éducation qui soulèvent la préoccupation sur la croissance des inégalités de bien-être et le ralentissement de la croissance économique. Le décrochage scolaire au premier cycle du secondaire est un problème important car il entraîne le manque de compétences de base chez les jeunes qui quittent l'école, qui auront des difficultés à trouver un emploi pour gagner de l'argent. Il mérite donc d'examiner plus avant ce qui se cache derrière la décision des enfants de quitter l'école. De plus, il n'y a aucune preuve sur la façon dont les compétences déterminent le retrait de l'école. Le phénomène n'a pas été étudié dans sa nature dynamique. Ainsi, un modèle de survie appliqué au décrochage tenant compte des compétences permettra de faire la lumière sur son rôle et sur d'autres facteurs de décrochage au fil du temps.

2) Comment le milieu social et l'environnement scolaire interagissent-ils pour déterminer les progrès de l'apprentissage?

Les enfants avec des désavantages sociaux sont plus susceptibles de faire face à de mauvais résultats scolaires. Dans ce cas, les politiques d'intervention fonctionnent si elles peuvent éliminer les obstacles à l'égalité des chances dans l'éducation. L'une de ces politiques est la mise en œuvre par le pays de l'enseignement à temps plein pour l'enseignement primaire. Alors que la scolarisation à temps plein s'est développée massivement depuis 2010, peu de preuves de son effet sur l'inégalité des chances sont disponibles. Par conséquent, l'étude de la relation servirait la politique fondée sur des preuves de la réforme du secteur de l'éducation. À cette fin, la recherche appliquée à la fois l'analyse descriptive et les modèles à valeur ajoutée en utilisant les données de l'enquête scolaire Young Lives en combinaison avec l'enquête auprès des ménages.

3) Comment les compétences prédisent-elles les résultats sur le marché du travail des jeunes?

À la suite des preuves sur la façon dont les compétences sont formées pour déterminer les résultats scolaires et l'inégalité, cette recherche vise à identifier l'importance relative des compétences pour les résultats sur le marché du travail. En abordant la question, cela éluciderait le rôle des compétences non cognitives sur les moyens de subsistance. Le paradoxe de la haute compétence cognitive et de la faible productivité du travail évoqués plus haut peut s'expliquer, donnant ainsi une indication de l'amélioration des compétences pour une productivité du travail plus élevée.

Ces questions de recherche sont traitées dans trois chapitres correspondants. Les résumés des trois chapitres sont fournis dans les sous-sections suivantes. Chaque résumé présente brièvement la motivation de la recherche, la méthodologie appliquée, les données et les résultats.

Chapitre 1: Les compétences comme prédicteurs du décrochage scolaire au Vietnam: une estimation du modèle de risque proportionnel de Cox

L'enseignement secondaire inférieur a été considéré depuis longtemps comme le niveau d'éducation fournissant des compétences de base pour l'apprentissage ultérieur ou le développement de la capacité pour les moyens de subsistance futurs. Pour cette raison, à l'instar d'autres pays et par l'engagement pour l'Education pour tous, l'enseignement primaire et secondaire inférieur est obligatoire au Vietnam. Cependant, seul le niveau primaire a été presque universel. En outre, les inscriptions dans le premier cycle du secondaire étaient d'environ 90 pour cent en 2013. Plus préoccupant, les taux d'abandon à ce niveau d'éducation sont restés à 11 pour cent depuis 2010. Les décrocheurs, sans compétences de base, seraient désavantagés pour obtenir des emplois décents, en particulier dans le contexte d'une intégration économique croissante. Par conséquent, il est devenu urgent que les décideurs comprennent pourquoi les adolescents abandonnent leurs études secondaires. Dans la littérature internationale sur les facteurs du décrochage, peu d'études considèrent l'aspect dynamique du phénomène. De telles preuves sont encore moins nombreuses pour le Vietnam, avec des limitations dans les approches et les données utilisées. Aucune étude ne prend en compte le rôle des compétences dans la détermination du phénomène, principalement en raison de l'indisponibilité de données appropriées. Par conséquent, ce chapitre étudie les facteurs de décrochage au niveau secondaire inférieur, en utilisant le modèle de risque proportionnel de Cox et les données de panel pour la cohorte plus âgée de Young Lives. Il identifie d'abord les types de compétences qui sont importants pour le décrochage, puis examine comment le pouvoir prédictif de ces compétences pour l'abandon varie au fil du temps.

Bien qu'ils ne traitent pas de l'effet causal des compétences, les modèles de Cox sont justifiés d'examiner l'abandon dans un contexte dynamique, en traitant le problème de la censure tout en tenant compte des facteurs variant dans le temps dans l'analyse. Le chapitre examine pour la première fois le phénomène dans sa

nature dynamique et donne la preuve sur le pouvoir prédictif des compétences pour le décrochage au fil du temps.

Les résultats de l'analyse montrent que les compétences cognitives, l'éducation des parents, les aspirations éducatives des adolescents et le temps consacré aux corvées ménagères sont les facteurs les plus significatifs du décrochage scolaire dans l'enseignement secondaire inférieur. Les effets des mathématiques, de l'aspiration éducative de l'adolescent et du statut socioéconomique (indice de richesse du ménage, taille du ménage et temps passé par les enfants aux tâches ménagères) sur la probabilité de décrochage à un âge plus précoce s'amplifient chez les plus jeunes.

Chapitre 2: L'enseignement à temps plein réduit-il les inégalités éducatives au Vietnam?

L'extension du temps scolaire pour atteindre le temps plein a été mise en place au Vietnam pour faire face au manque de temps d'enseignement dans l'enseignement primaire en comparaison des normes internationales. L'allongement du temps d'enseignement devrait améliorer les résultats d'apprentissage. La scolarisation à temps plein peut avoir un impact sur l'inégalité en matière d'éducation en comblant le manque de temps d'enseignement entre enfants viennent des familles de différentes conditions. En effet, les familles aisées utilisent les cours privés pour leurs enfants, ce qui pourrait augmenter les inégalités éducatives. Il y a peu de preuves sur l'impact de la scolarisation à temps plein sur l'inégalité des chances, . Par conséquent, ce chapitre examine si le développement massif de la scolarisation à temps plein au cours des dernières années peut améliorer l'égalité des chances au Vietnam en utilisant les données de l'enquête école Young Lives 2011-2012. Tout d'abord, il examine la variation du temps d'enseignement et des ressources scolaires entre les élèves avec et sans scolarisation à plein temps de différent milieu sociaux. Il étudie ensuite les progrès d'apprentissage associés à la scolarisation à temps selon les différents milieux sociaux par l'estimation de modèles à valeur ajoutée. Pour faire face à la surreprésentation des enfants de Young Lives dans l'enquête sur les écoles, ces modèles à valeur ajoutée incluent les rapports inverses de Mills. Les ratios sont obtenus à partir du modèle de sélection en deux étapes de Heckman estimé de la probabilité pour les enfants de Young Lives enquêtés en 2002 d'être inclus dans l'enquête écoles. L'estimation des modèles pour des sous-groupes de milieu sociaux est également menée pour étudier l'hétérogénéité de l'effet de la scolarisation à temps plein. Les données suggèrent que la scolarisation à temps plein ne peut pas réduire l'inégalité des chances dans la réussite de l'apprentissage à ce stade de la mise en œuvre. Au contraire, il contribue à augmenter le fossé dans l'apprentissage des progrès entre les enfants de différentes positions sociales.

Chapitre 3: Les compétences cognitives ou non cognitives sont-elles importantes pour prédire les gains du marché du travail?

Bien qu'il existe un consensus sur le pouvoir prédictif des compétences cognitives et non cognitives dans l'explication des gains ou des salaires, l'importance relative de ces compétences est encore discutée. Contrairement aux abondantes études qui fournissent des preuves mixtes sur les effets de ces compétences pour les pays développés, peu de preuves existent pour les pays en développement. Aucune étude de ce type n'est disponible pour le Vietnam. En outre, étant donné le haut niveau de compétences cognitives mais de faibles compétences non cognitives au Vietnam, identifier quel type de compétences est plus important fournit un éclairage importante pour des politiques visant à augmenter les compétences pour améliorer les résultats du travail. Ce chapitre utilise les données du panel Young Lives pour examiner la relation entre les deux types de compétences et les résultats du travail, y compris l'emploi et les gains.

Le chapitre applique le modèle d'équations structurelles, qui permet aux compétences cognitives et non cognitives, ainsi qu'aux caractéristiques individuelles, familiales et communautaires, de déterminer les mesures de compétences observées et les choix de la personne en matière d'éducation et de travail. Les résultats de la régression montrent un effet plus élevé de la capacité non cognitive que cognitive sur l'emploi et les gains. De l'échantillon, les personnes peu qualifiées sont plus susceptibles d'être employées, avec un pouvoir prédictif plus élevé de la compétence cognitive. Parmi ceux qui ont un emploi, ceux qui ont des compétences cognitives plus élevées ont tendance à travailler comme travailleurs indépendants, tandis que les plus qualifiées ont tendance à être des travailleurs salariés et ont souvent des revenus plus élevés.

Résumé

L'éducation a connu des progrès remarquables au Vietnam au cours des deux dernières décennies. Cependant, l'inégalité des chances en matière d'éducation aggrave les disparités chez les enfants et menace les progrès en termes de productivité du travail. Pourtant, les mécanismes de transmission des inégalités et le rôle des compétences dans ce processus restent encore largement méconnus. L'objectif de cette thèse est d'apporter un nouvel éclairage sur ces questions. Plus précisément, elle vise à étudier comment les compétences interagissent avec les facteurs environnementaux dans l'atteinte des résultats scolaires et l'insertion sur le marché du travail au Vietnam. Le premier chapitre de cette thèse examine dans quelle mesure les compétences prédisent l'abandon scolaire. Le chapitre 2 analyse l'effet du passage au temps complet d'enseignement au primaire sur les inégalités scolaires. Enfin, le chapitre 3 étudie l'importance relative des compétences dans l'insertion des jeunes sur le marché du travail. À cette fin, diverses approches quantitatives sont menées à partir des données de Young Lives. Les contributions de la thèse à la littérature existante sont de prendre en compte les compétences non cognitives dans l'analyse, de considérer l'interaction entre le milieu social et l'environnement scolaire et enfin de traiter les erreurs de mesures des compétences au Vietnam. La thèse montre que les efforts en vue de plus d'égalité d'opportunités scolaires doivent être poursuivis. Elle montre également l'importance d'améliorer les compétences non cognitives pour améliorer le bien-être individuel et la croissance économique.

Mots Clés

Compétences, inégalité des chances, abandon scolaire, extension du temps scolaire, progression des apprentissages, rendements des compétences, Vietnam

Abstract

Vietnam has attained outstanding performance in education during the past two decades. However, inequality in educational opportunities aggravates disparities among children and threaten the improvement of labour productivity. The underlying mechanism for transmitting inequality and role of skills in the process remaining unclear motivates this thesis. It aims to study how skills interact with environmental factors to determine outcomes on education and livelihoods in Vietnam. Firstly, Chapter 1 examines how skills predict dropping out. Then Chapter 2 inspects whether full-day schooling reduces educational inequality. Finally, Chapter 3 investigates the relative importance of skills in determining labour market outcomes. To this end, the research applies diversified quantitative approaches using data from the Young Lives in Vietnam. The contributions of the thesis are threefold, namely taking into consideration the non-cognitive skill in the analysis, addressing the interaction between social background and school environment, and the measurement errors accompanied poor proxies for skills in Vietnam. The thesis hints at further levelling educational opportunities of children and enhancement of non-cognitive skills for greater outcomes in life and economic growth.

Keywords

Skills, inequality of opportunity, dropping out, full-day schooling, learning progress, returns to skills, Vietnam