Education Dynamics in a Developing Country:Evidence from Indonesia

A Dissertation presented

by

Elif Deniz Gulenc Sumengen

to

The Graduate School

in Partial Fulfillment of the

Requirements

for the Degree of

Doctor of Philosophy

in

Economics

Stony Brook University

May 2016

Stony Brook University

The Graduate School

Elif Deniz Gulenc Sumengen

We, the dissertation committe for the above candidate for the

Doctor of Philosophy degree, hereby recommend

acceptance of this dissertation

Mark R. Montgomery - Dissertation Advisor
Professor of Economics

Warren Sanderson - Chairperson of Defense
Professor of Economics

Thomas Muench - Committee Member
Professor of Economics

Rebekah Burroway - Outside Member
Assistant Professor of Sociology

This dissertation is accepted by the Graduate School

Charles Taber
Dean of the Graduate School

Abstract of the Dissertation

Education Dynamics in a Developing Country:Evidence from Indonesia

by

Elif Deniz Gulenc Sumengen

Doctor of Philosophy

in

Economics

Stony Brook University

2016

There is an urgent need for upper secondary-level and above educated people in Indonesia. According to a recent report, Indonesian companies cannot fill 50% of their entry level positions. To increase the educational attainment, government has been implementing various polices, such as school construction program, compulsory education and allocating 20% of its budget to education, but inadequate enrollment at upper secondary and tertiary levels and the quality of education still remain as big problems. To shed a light on these urgent and recent problems, I ask four questions: a) What is the effect of school quality on tertiary attainment? b) Which school level does the quality matter more? c) Which factors prevent high ability individuals to get tertiary attainment? d) How important is parental background for educational attainment at each stage of the educational path? Previous work focused on school quality's effect on lower secondary education, and a lack of upper secondary-level and above educated people is an issue only brought up recently and analyzed in this paper. I show that primary school quality has a direct effect on tertiary attainment besides its indirect effect due to the accumulation of school quality at each level. To generate my dataset, I use four waves of Indonesia Family Life Survey. My model accounts for unobserved heterogeneity to handle self selection issues in education. This is one of the few studies in a developing country modeling long term educational decisions. I analyze the role of family background, location, personal characteristics, number of schools used in each community, primary school quality, as well as student's ability and motivation for transitions to lower secondary, upper secondary, and tertiary education in Indonesia. With a focus on tertiary educational attainment, I show that long term factors and early fundamental education play a big role. These findings further support the importance of promoting cognitive ability and high quality education early in life; especially for those who are coming from more disadvantaged environments.

Dedication Page

For my grandmother Elif.

Growing up, I listened to several stories from my parents about how difficult it was for them to even attend a primary school. They would walk at least an hour from their village before they could finally reach the nearest town where their primary school resided. Both of my parents were born in a rural part of Turkey, where in the 1960s even finding drinkable water was hard to come by yet alone a school. Despite the circumstances, my mother was still among the fortunate; my beloved grandmother, who gave birth to 6 sons and 5 daughters including my mom, believed in having a good education was essential. Perhaps because she herself did not know how to read and write, she insisted that all her kids were educated equally, regardless of their gender. My grandmother has always been a hero and an inspiration for me to pursue a study on “Education Dynamics in a Developing Country”. I am doing my best to be a grand-daughter that you would be proud of if you were alive.

Contents

List of Figures

List of Tables

Acknowledgements

First, I would like to thank all my committee members. I would like to thank my advisor Prof Mark R. Montgomery for his guidance and help. I started my journey with networks and migration, and then I ended up writing on human capital. I learned a lot while exploring for my PhD, and my advisor encouraged me to search and find my own way. He helped me to gain a real research experience. I would also like to thank my co-chair Prof Warren Sanderson. He always helped me to see the big picture. His vision and questions strived me to explore and enjoy what I am doing more. He made me love economics even more. I would like to thank my other committee member Prof Thomas Muench for his deep questions. He pushed me to understand the equations and the meaning of each equation in detail. He taught me to question even the basic assumptions made by Econometricians. I would like to thank my outside member Prof Rebekah Burroway. She gave me feedback from the sociological perspective, which helped me to create an interdisciplinary thesis. I would also like to thank Samuel Centorrino, Hugo Benitez-Silva, and other members of the Stony Brook University Applied Micro Workshop for their comments and suggestions.

Second, I would like to thank my friends: Muge Kocak, Ozge Yucel, Aram Kim, Abhinav Anand, Bruno Badia, Qian Li, Biligbataar Tumendemberel, Lunan Jiang, Aylin Celtik, Selin Gonen, Anju Bhandari, and Tamara Gurevich for their support and their friendship.

Third, I would like to thank my beloved mother Cennet Akten-Gulenc and my father Halil Gulenc for their tremendous support. I would also like to thank my parents-in-law, Havva and Dogan, and my sisters-in-law, Demet and Dilek for listening to my presentations whenever I need them to. I thank my dear cousin Bikem Akten for listening to my complaints and taking care of my daughter whenever I need her.

Last, I would like to thank my husband, Baris Sumengen and my daughter, Ada Sumengen for their continuous love and support.

Chapter 1

Motivation

Human capital is one of the crucial factors for economic progress. Most of the developing countries face shortage of well-educated people, which causes a lower level of labour productivity. One of these developing countries is Indonesia, which has the world's fourth-largest education system with roughly 55m students, 3m teachers, and more than 236,000 schools in 500 districts.[1] A recent report prepared by the Boston Consulting Group (BCG) states that due to the lack of educated people at tertiary levels, Indonesian companies face the problem of filling half their entry level positions and this number is expected to increase to more than 70% by 2025.[2] According to World Bank Report, there are about 19 million net new jobs in Indonesia since 2002, of these 6.1 million to tertiary graduates (WorldBank, 2014). See Table 1.1 for

[1] http://www.economist.com/news/asia/21636098-indonesias-schools-are-lousy-new-administration-wants-fix-them-schools

[2] http://wenr.wes.org/2014/04/education-in-indonesia/

educational attainment in IFLS surveyed areas of Indonesia.[3]

My objective in this dissertation is to analyze what affects the educational attainment, with a specific focus on tertiary attainment in a developing country. First, I would like to analyze the key factors affecting schooling decisions at each stage of the educational path, i.e. primary, lower secondary, upper secondary, and tertiary, since it is important to understand how the education inequities take place from the start. Second, I would like to unearth some casual relationships that could lead policies to encourage more high ability individuals to fulfill their potential and get university degrees.

Table 1.1: Education attainment in percentage for people born between 1978 and 1984 (IFLS surveyed areas)

No Education	Primary	Lower Secondary	Upper Secondary	Tertiary
0.00	0.29	0.27	0.34	0.09

Education is a central concern for the Indonesian government. Indonesia has implemented various policies to increase educational attainment over the last few decades and invested a huge portion of its budget in education. Here are three important policy implementations:

First, in the 1970s primary school construction was a focus and 61,000 primary schools were built under INPRES program (Duflo, 2001). Figure 1.1 shows that almost all of the people born in the 70s and later have attended

[3]The Indonesian Family Life Survey (IFLS) is an on-going longitudinal survey in Indonesia and a representative of about 83% of the Indonesian population. Individuals who dropped earlier than graduation are included in corresponding levels.

some level of school. This demonstrates the success of INPRES program (implemented in the 1970s) and the 6-year compulsory education (implemented in 1984).

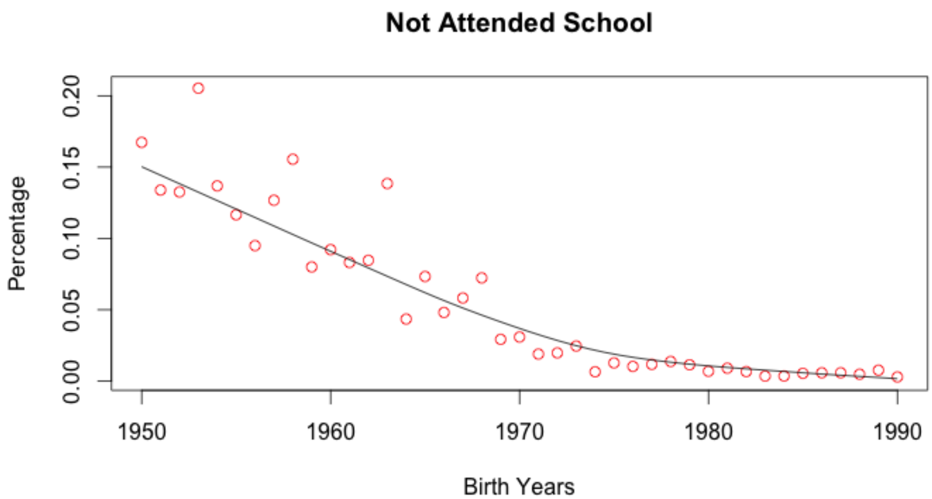


Figure 1.1: Percentage of individuals who have not attended school in IFLS surveyed areas

Second, Indonesia fully implemented six year compulsory education in 1984. Over the next 10 years, they slowly worked on launching nine year compulsory education and in 1994 the new compulsory law went into effect.[4] Nine year compulsory education was planned for complete implementation by the end of 2003/2004.

Third, to encourage lower secondary education, during the 1980s a large

[4]http://education.stateuniversity.com/pages/662/Indonesia-EDUCATIONAL-SYSTEM-OVERVIEW.html

number of lower secondary schools were constructed and fee subsidies were offered after 1994. Construction of lower secondary schools accelerated during the 1980s. Out of 171 rural villages surveyed in IFLS, only 95 of them had a lower secondary school by 1980. During the 80s, 47 more villages had their first lower secondary school constructed. Moreover, for those students who drop out of primary or lower secondary school, the Ministerial Decree 0131/U/1994 was released to provide out–of–school education programs equivalent to primary and lower secondary school.

Even though Indonesia has boosted primary and lower-secondary enrollment rates with all these policy interventions, there remains inadequate enrollment at upper secondary and tertiary levels. Improving educational quality has been difficult for Indonesia as well. Indonesian students rank the second lowest in PISA 2012 (Program for International Student Assessment) out of 65 countries (OECD, 2012).[5] This highlights the importance of focusing on both quantity and quality of education in Indonesia.

To understand these education dynamics of Indonesia, I use the Indonesian Family Life Survey.[6] I took a sample that consists of 1,776 students born between 1978 and 1984. This cohort is chosen so that when they start school, the INPRES program and compulsory primary education is fully in effect. My latest data are from the IFLS 2007 wave. Thus, I took birth years of 1984 latest so that they have a chance to finish their education fully by

[5]PISA is an assesment test that compares the performance of 15 year-olds in 65 countries in reading, mathematics and science every three years.

[6]http://www.rand.org/labor/FLS/IFLS.html

2007.

Since I try to analyze the key factors at each stage, a sequential choice model (see Figure 5.1) is used to model individuals schooling decisions. Note that there is a self selection and selection bias problem in education research, which needs to be addressed. Motivation and ability of students are usually not observed and high ability (motivation) people may find school less difficult and be more likely to progress. As grades increase, low ability students are likelier to drop out than high ability students. To model this process, I incorporate unobserved factors into the model as unobserved heterogeneity. This unobserved heterogeneity connects decision process in all educational stages, and needs to be integrated out (marginalized) using its distribution. The advantage of this approach is to model the selection bias in schooling considering both observables and unobservables with the possibility of deducing the preferences of individuals. More specifically, I attempt to analyze factors that help to increase the educational attainment of the high ability students.

After controlling for unobserved heterogeneity, I show that parental background, region of residence, number of schools, and child's cognitive skills and non-cognitive personality traits are key factors of educational path decisions. Household per capita consumption, father's education, and mother's education are important in all stages of educational choices. Living in a rural areas as opposed to an urban area negatively affects the earlier stages of educational path but does not seem to have an effect in the last stage. Being female has a negative effect on lower secondary school decision, and has a positive

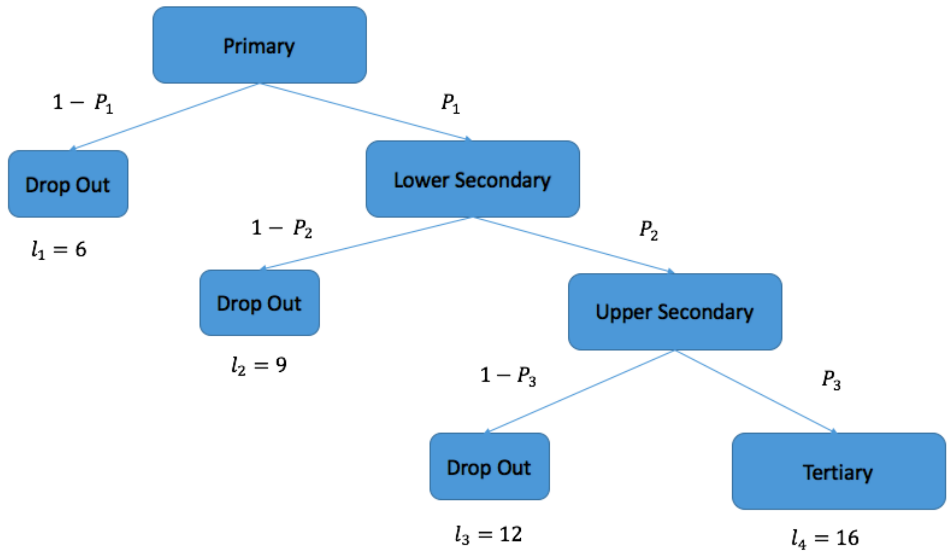


Figure 1.2: Educational Attainment Diagram

effect on tertiary decision. Birth order has a positive effect for second stage. Innate ability affects all stages of the educational path.[7] Non-cognitive personality trait, i.e. smoking before age 15, has a negative effect for all stages.[8] A striking result is that primary math scores have a direct effect on whether to continue to tertiary education or not.

These findings further support the importance of promoting cognitive ability and high quality education early in life. To investigate this further, I

[7]I pick Raven matrices questions from IFLS administered data as a representative of innate ability. Math questions on the other hand reflect the learnings of people during their school years and used as a representative of learned ability.

[8]Following the previous literature (Heckman and Mosso, 2014), I picked smoking before age 15 as a proxy for non-cognitive personality traits. In a 99% Muslim country with highly religious and conservative societal norms, it is difficult to quantitatively track other risky behaviors among those younger than 15 such as regular drinking or intercourse.

create a school quality index using test scores from students randomly sampled from each school. I show that primary school quality is most important compared to other levels and has a long term effect in child's education.

The results found so far suggest that long-term factors may have a stronger influence on the determinants of schooling attainment. Previous literature, Cameron and Heckman (1998), also point to the importance of long-term factors as opposed to short-term credit constraints.

The main contributions of this paper could be summarized as follows:

- This is one of the few studies on schooling transitions in a developing country addressing self-selection issues. I separate and analyze the effects of different ability types on schooling decisions. To the best to my knowledge, there are not many previous studies showing the effect of family background and abilities on transition to different stages of the educational path in Indonesia. My study shows the importance of various interventions during different periods of an individual's educational life cycle. The government of Indonesia aims to promote equal education opportunities for all children and also to increase the tertiary enrollment rate. Therefore, it is crucial to understand the important factors for transition to lower secondary and upper secondary education besides parental background.

- IFLS is a rich and large dataset, which let me to study the importance of innate cognitive ability, learned cognitive ability, and non-cognitive personality trait on educational attainment. This study identifies different effects of cognitive ability and non-cognitive personality traits across an individual's schooling transition in a developing country. Besides cognitive ability and non-cognitive personality traits, there is also information on the EBTANAS (National Achievement Test) math scores for a random sample of 25 students for each school surveyed. Using this information, this research also attempts to understand the effects of school quality on educational attainment in Indonesia.

The rest of the paper is organized as follows. Chapter 2 covers the background of the Indonesian education system. In Section 2.1 history of the education system in Indonesia is given. Then, in Section 2.2 education in Indonesia is compared with other countries. Chapter 3 reviews previous research on education. First, in Section 3.1, previous research on education in developing countries is summarized. Then, Section 3.2 reviews the usage of sequential choice models in the education literature. Last, Section 3.3 summarizes the literature with a special focus on abilities. Chapter 4 introduces the details of data, covariates, and descriptive statistics. In Section 4.1, the specifications of Indonesian Family Life Survey explained. A detailed explanation of how I picked the cohort using IFLS is outlined in Section 4.2. In Section 4.3, important descriptive statistics are summarized. In Section 4.4, I explain the covariates picked for the model. Chapter 5 shows the model and results. In Section 5.1, sequential choice model is described. In Section 5.2, results are presented. Chapter 6 introduces the policy simulations. Finally, Chapter 7 gives the conclusion and the implications for future work.

Chapter 2

Education in Indonesia

In Section 2.1, history of the education system in Indonesia is given. Then, in Section 2.2, education in Indonesia is compared with other countries.

2.1 The History of Education System in Indonesia

Indonesia's educational system is still in the early stages of its development. During the colonial era, primary education was first introduced by the Dutch in Indonesia and schooling was reserved for the Europeans only. Towards the end of 19th century, the system was opened to the Indonesian aristocracy as well. The right of every Indonesian citizen to get an education was declared in 1950 after the Indonesian Independence in 1949. Initial expansion of schools and education system were slowed down under the regime of General

Sukarno. In 1965, General Sukarno was replaced by Suharto.

The big change in education happened in the early 1970s. In order to increase the level of education, the Indonesian government initiated compulsory education and started a school construction program (INPRES) in 1973 with the proceeds from high oil prices. Between 1973 and 1978, the government built 61,000 primary schools, an average of 2 schools per 1,000 children of primary school age. The government of Indonesia allocated more schools in provinces where the initial enrollment rate was low. In 1978, the enrollment rate reached 84 percent for males and 82 percent for females (WorldBank, 1989). Duflo (2001) shows that the INPRES program led to an average increase of 0.25 to 0.40 years of education (0.12 to 0.19 years for each new school built per 1,000 children) and a 3 to 5.4 percent increase in wages.

Indonesia fully implemented six-year compulsory education in 1984. Then 10 years later, they launched nine-year compulsory education in 1994.[1] Nine year compulsory education was planned for completion by the end of 2003/2004 with an intention to enter the global market: AFTA (Asia Free Trade Area) in 2003 and APEC in 2010.[2] Up until 1994, the emphasis of the nine-year compulsory education was put on enrolling all children of 13–15 years of age in lower secondary schools. Many factors could prevent primary school graduates from continuing to lower secondary schools, perhaps the most important being the low economic status of their family. The Ministry of Ed-

[1]http://education.stateuniversity.com/pages/662/Indonesia-EDUCATIONAL-SYSTEM-OVERVIEW.html

[2]http://unesdoc.unesco.org/images/0014/001470/147087e.pdf

ucation issued a Ministerial Decree (No. 0151/K/1994) to eliminate lower secondary level fees and Ministerial Decree (No.0131/U/1994) to provide out-of-school education programs for those who drop out during primary and lower secondary education. On top of that, The National Foster Parent Family provided fellowships to students with economic difficulty. Unfortunately, in 1998, the Asian Financial Crisis hit Indonesia and GDP fell by 12% that year. Thomas et al. (2004) found that poor households cut educational expenses for younger children in order to provide education for older ones during the crisis. Cameron (2002) used 100 Villages Survey and concentrated on rural areas of Indonesia, finding that enrollment rates dropped slightly during the crisis and rebounded higher than pre-crisis levels later. Strauss et al. (2004) found similar results showing a mid-term recovery of the crisis.

Another implementation in the education system in Indonesia was decentralization. Before 1999, the education system of Indonesia was centralized. After the enactment of Law No. 22 of 1999, local governments have become responsible for the provision of basic services. The Law 22/1999 broadly outlines powers and responsibilities of each government level and abolishes any hierarchal relationship between districts/municipals.

Since 2009 Indonesia allocated a fifth of its annual budget to education. There is a high enrollment rate in primary school, reaching 97 per cent in 2009, but there is still not enough enrollment at the upper secondary and tertiary levels. The government of Indonesia has been planning to implement

12 years of compulsory education starting by the end of 2015.[3]

Improving educational quality has been another challenge for Indonesia besides enrollment problem at the upper secondary and tertiary levels. Even though Indonesia allocates 20 percent of its annual budget to education, Indonesian students rank the second lowest in PISA 2012 out of 65 countries (See Figure 2.1) and 38 out of 42 countries in TIMSS 2011 (Mullis et al., 2011).[4]

2.2 Comparing Education in Indonesia to Other Countries

Based on UNDP Report 2014 on Human Development Index that measures achievement in terms of life expectancy, educational attainment and adjusted real income, Indonesia is located in number 110 out of 188 countries surveyed. From East Asia and the Pacific, countries which are close to Indonesia in terms of 2014 HDI rank and to some extent in population size are the Philippines and China, with HDIs ranks 115 and 90 respectively.

When we look at World Bank data on gross primary, secondary and tertiary enrollment ratios in years 1995 and 2000, we could see the significant

[3] http://www.thejakartapost.com/news/2015/01/13/12-year-compulsory-education-start-next-june-minister.html

[4] The Trends in International Mathematics and Science Study (TIMSS) is a series of international assessments of the mathematics and science knowledge of students around the world.

Figure 2.1: Pisa Test 2012 results (Source: OECD Report)

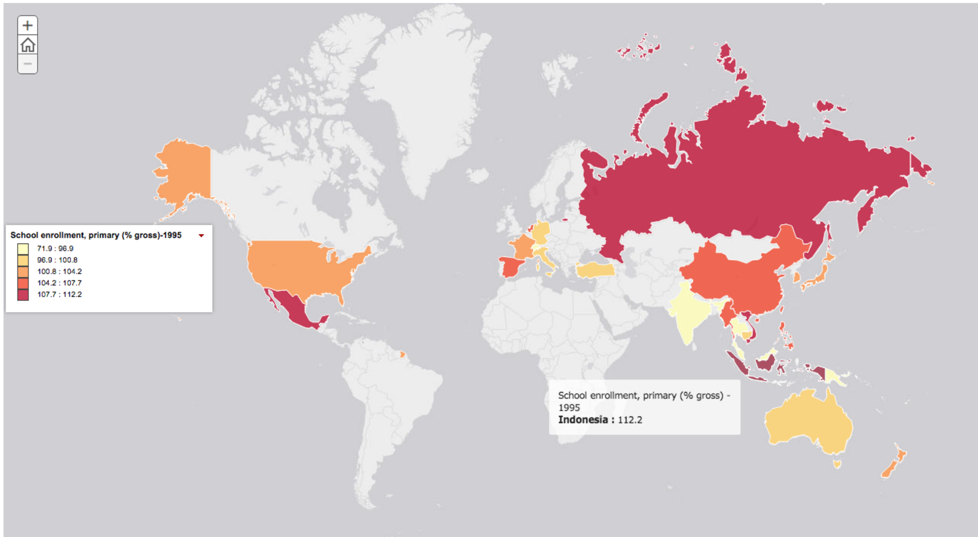


Figure 2.2: Gross Primary Enrollment Rate

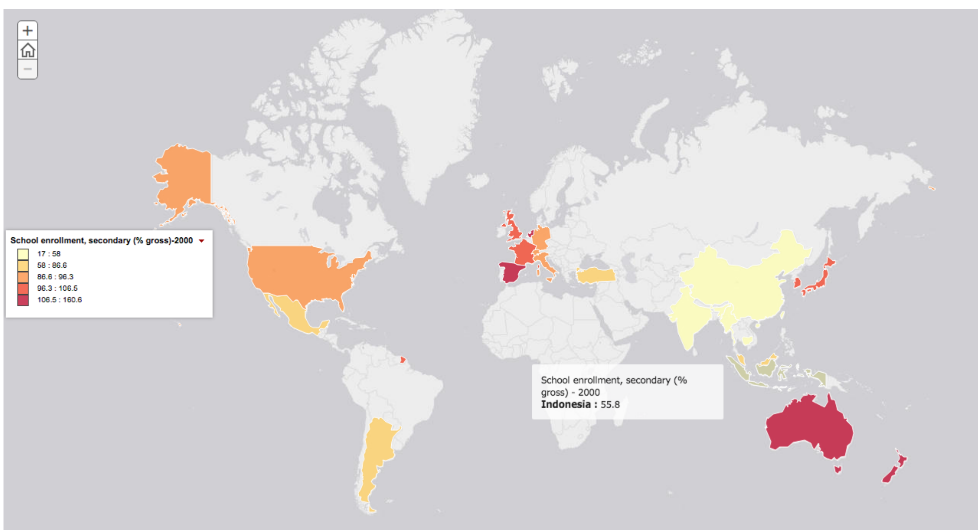


Figure 2.3: Gross Secondary Enrollment Rate

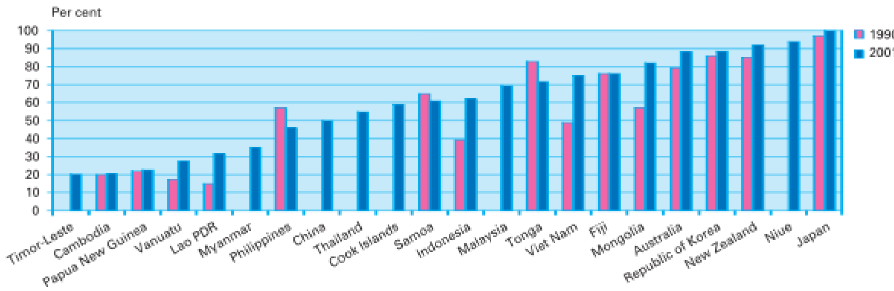


Figure 2.4: Net enrollment Ratio in Secondary Education

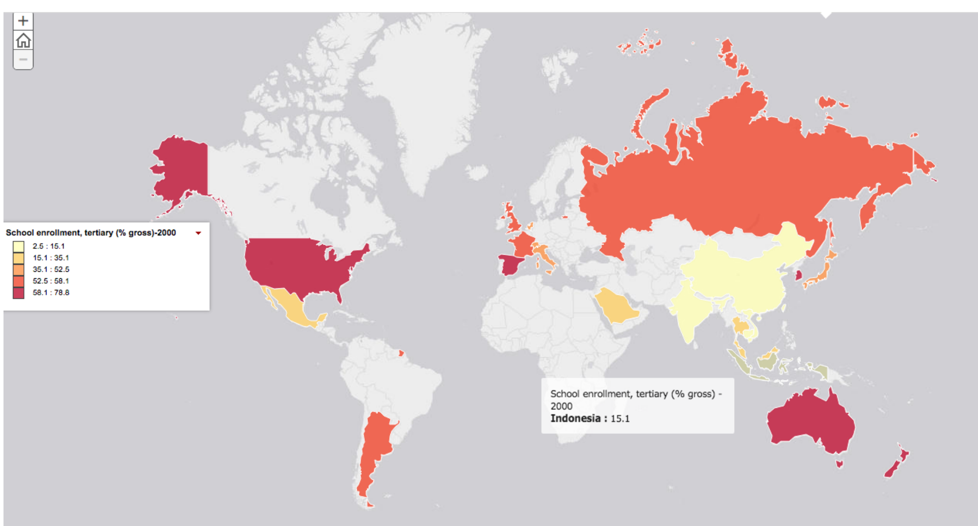


Figure 2.5: Gross Tertiary Enrollment Rate

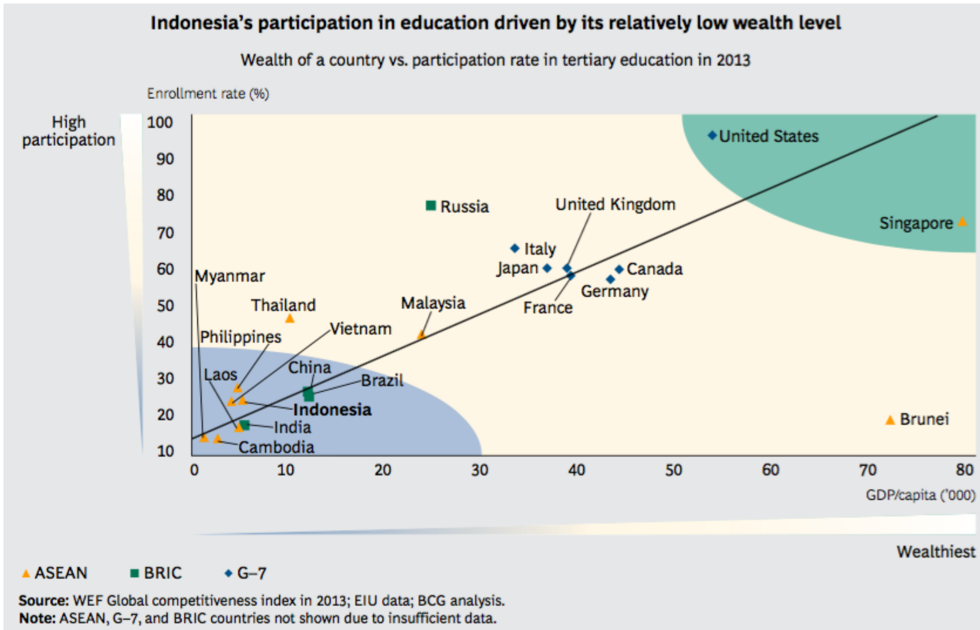


Figure 2.6: Tertiary Enrollment Rate (Source: BCG Report)

success of Indonesia on primary enrollment ratio in Figure 2.2. Indonesia had an average of 112% gross primary school enrollment rate in 1995, which is higher than developed countries such as Australia and the United States. Gross enrollment ratios are computed as the total number of students enrolled in the corresponding level as a percentage of school-age population adopted by UNESCO. Primary school age is 7–11, secondary school-age is 12–17 and tertiary school-age is 20–24 for most of the countries. Because of the way this indicator is constructed, we could see the ratios over 100% in any schooling level as long as enough people from outside the corresponding age group enrolled in that level.

The gross secondary enrollment rate reached almost 55.8% in Indonesia in 2000. The net enrollment rate is higher than secondary enrollment rates observed in China but lower than the neighbor Malaysia as can be seen in Figure 2.4. Indonesia's gross secondary enrollment rate is considerably lower than those of Australia and New Zealand as can be seen in Figure 2.3.

The rate of return to tertiary education is really low in Indonesia as can be seen in Figure 2.6 and Figure 2.5 with 15.1% behind the neighbor Malaysia with 25.7% and higher than India with 9.5% and China with 7.8%.

In developing countries, the proportion of people receiving secondary and tertiary education is low compared to more developed countries. The question is why? Two potential reasons behind this are credit constraints and parents background. Indonesia is a good representative of ASEAN countries. The key findings of this paper for Indonesia might be applied to other ASEAN countries in the region.

Chapter 3

Previous Work on Educational Attainment

First, in Section 3.1, previous research on education in developing countries are summarized. Then, Section 3.2 reviews the usage of ordered choice models in the education literature. Last, Section 3.3 summarizes the literature with a special focus on abilities.

3.1 Education in Developing Countries

The education dynamics in developed and developing countries are different, mainly because of the huge gap in education quality and access. Buchmann and Hannum (2001) examined studies published till 2001 on education and inequality in developing regions. Glewwe et al. (2011) reviewed the research between 1990 and 2010 to investigate which specific school and teacher characteristics appear to have strong positive impacts on learning and time in school. Using these two extensive reviews and recent studies in the literature, this section summarizes the education in developing countries under three titles: a) family background and educational outcomes, b) the quality of teacher and school effects, and c) state policies for educational attainment. This section also includes studies from four different fields including sociology, education, psychology, and economics literature.

3.1.1 Family Background and Educational Outcomes

Checchi (2006) summarized how educational choices of future generation affected by previous generation with four channels. One channel through the transmission of talent. This can be genetic, such as race, beauty, and height. Using the empirical evidence from the samples of twins, Bowles and Gintis (2002) found that IQ is not a major contributor to the inheritance of economic status whereas genetic transmission of earnings and enhancing traits appears to play a role. The other channel is through cultural influences. This means children with more educated parents will be more educated. A third channel of intergenerational persistence comes from liquidity constraints. From an empirical point of view, it is hard to distinguish cultural links to financial links since both of them are correlated. However with the richest population who are not liquidity constrained, there is at least a possibility to get an approximate estimate of it. A fourth channel comes from territorial segregation and it is related to family wealth. If residential choices are influenced by the evaluation of local school quality, then school quality might affect house prices and this affects the wealth of the people in that neighborhood.

There are varying studies in developing countries focusing on the effect of the family background to schooling progress. Coleman et al. (1966) first showed that family background was more important than school factors in the U.S. After that, Heyneman (1976) have done similar research on a developing country and presented the opposite in Uganda. Following his work, Heyneman and Loxley (1983), Fuller (1987), and Fuller and Clarke (1994) found similar results showing family background to be less important than school factors in developing countries. Their conclusion is that the poorer the country, the greater the impact of the school quality on educational attainment.

In contrast to the previous research using production function, recently using multilevel analysis, studies found greater effect of family background than school effects in developing countries such as in Zimbabwe (Riddell, 1989) and Thailand (Lockheed and Longford, 1991). Lillard and Willis (1994) also did a related study and used Malaysian data to study the family decisions about the schooling transitions of individuals. They allowed for siblings correlations. They found that at least two-thirds of the impact of parental education on childrens schooling is a direct consequence of parent schooling while the remaining one-third can be attributed to unmeasured factors that influence educational attainment of parents and children. Pal (2004) showed that the same set of individual/parental/household characteristics may affect different levels of schooling differently in Peru. He found that especially at primary/secondary level crucially depends on whether children have educated parents and more siblings in the working age category.

Research has also examined how family structure and the number of children affect the educational attainment in developing countries. Parents face trade-off while making decisions for their children's educational attainment regarding the size of their family (Becker and Lewis, 1973). Studies on developing countries has different results on sibship size. Research in the U.S. shows that there is an inverse relationship between number of sibling and educational attainment. Parish and Willis (1993) showed that later born sibling gets more education than early born siblings in Taiwan. Montgomery and Lloyd (1997) also found no effect of excess fertility on educational progress in the Philippines. Buchmann (2000) found no impact of sibship size on educational attainment in Kenya. Maralani (2008) mentioned that for developed countries there is a negative correlation between family size and children's schooling, but for developing countries this correlation ranges from positive to negative depending on the context. She analyzed the effect of number of children for Indonesia and showed that in urban areas, the correlation between family size and childrens schooling was positive for older cohorts but negative for more recent cohorts. She also emphasized that in rural areas there is no significant association between family size and childrens schooling for any cohort.

3.1.2 The Quality of Teacher and School Effects

In this section, I first focus on the school quality and math education. Then I summarize the recent review done by Glewwe et al. (2011) on the impact of school and teacher characteristics on learning.

Math Education and School Quality

Hanushek et al. (2008) used panel data on primary school age children in Egypt to estimate the relationship between school leaving behavior to school quality. They found that, holding student's ability and achievement constant, a student is less likely to continue schooling if he attends a low quality school. This rational behavior suggests that common arguments about a trade-off between quality and quantity of schools may misrepresent the problem and give rise to limited investment in school quality.

A low-quality school may leave a student behind to learn the skills needed for the next grade level, and cause psychological costs to get more education (Rouse and Barrow, 2006). Thus, children who start ahead may stay ahead, and children who start behind may stay behind.

Duncan et al. (2007) showed that there is a substantial correlations between early and later knowledge and especially differences among children in mathematics knowledge, which is even more stable than in reading and other areas. Siegler et al. (2012) pointed that –after controlling for general intellectual ability, working memory, family income and education– primary school knowledge of fractions and division predicts the upper secondary school mathematics achievement. They showed in their paper that the correlation between primary school fractions knowledge and high school mathematics achievement was expected, but the relation was not, and pointed out the importance of teaching mathematics properly.

School and Teacher Characteristics

Glewwe et al. (2011) summarized 79 studies from developing countries with a focus on which specific school and teacher characteristics appear to have strong positive impacts on learning and time in school under three sub-titles: a) pedagogical materials and school infrastructure, b) teacher and principal characteristics, and c) school organization.

- Pedagogical materials and school infrastructure: Intuitively, it is natural to think that both pedagogical materials and school infrastructure has a positive effect on student learning. After eliminating less rigorous studies, Glewwe et al. (2009) showed that there is a weak evidence in the literature for textbooks and exercise books to be beneficial for student learning. The findings for blackboards and other visual aids are generally positive, and the quality of the schools walls, roofs, desks, tables and chairs offered strong support that improvements in these school characteristics raised students test score.

- Teacher and principal characteristics: Teachers' knowledge of the subjects that they teach shows very strong positive effects (Metzler and Woessmann, 2010), whereas teacher's education and experience show weak effects on students learning. Tutoring has also been examined by a randomized trial, the study of the BALSAKHI tutoring program in India by Banerjee et al. (2007). The study showed providing tutoring to students falling behind in the curriculum increased their test scores.

- School organization: School organization effect could be summarized as follow. First, crowded class size usually decreases student learning, as one would expect, but this is not always the case (Urquiola, 2006). Second, the results for teacher absenteeism are clearly negative (Suryadarma et al., 2006). Last, the hours of the school day and multi-grade classrooms have unambiguous results with respect to previous studies.

3.1.3 State Policies for Educational Attainment

A significant research covers how state policies, including compulsory schooling, fee subsidies, and school expansions affect education. King and Lillard (1987) examined how individuals' family background and government educational policies together influence schooling levels, and found that policies have significantly affected attainment and distribution levels in both countries, the Philippines and Malaysia. Eckstein and Zilcha (1994) showed that providing compulsory schooling increases the economic growth and in the long run also helps majority of individuals in each generation to be better off. Duflo (2001) showed that the schooling and labor market consequences of school construction program in Indonesia. She mentioned that the INPRES program (school construction program) led an average increase of 0.25 to 0.40 years of education (0.12 to 0.19 years for each new school built per 1,000 children) and led to 3 to 5.4 percent increase in wages, which implies estimates of economic returns of education ranging from 6.8 to 10.6 percent.

Education is critical to promote economic well-being in developing countries. International institutions, such as UNICEF, UNESCO, and the World Bank, have also pursued the expansion of schooling as a crucial component of development. These institutions started The Education for All (EFA) movement to provide early childhood education, gender equity, and quality basic education for all children, youth and adults. Government policies and the help of the international institutions to provide the education equality is very important especially for developing countries to catch up the education level of developed countries.

3.2 Sequential Choice Models in Education

Previous work in the literature on sequential choice (grade progression model) focused mostly on developed countries. Mare (1980) first showed the importance of sequential choice in the education system and decomposed the final educational attainment into a series of stages in sociology literature. The sequence of grade transition probabilities constructs the probability of schooling attainment and dividing schooling into stages provides in detail analysis in schooling progress. Most crucially, educational selectivity is also considered in grade progression model by Cameron and Heckman (1998), Cameron and Heckman (2001). Since low ability people drop out more compared to high ability people at early stages of education. Considering only upper secondary school graduates or lower secondary school graduates will create a selection bias in educational attainment analysis. Cameron and Heckman (2001) and Cameron and Heckman (1998) used grade progression model and found that the long-term factors such as parental educational background and child ability to be key factors of continued schooling. Cameron and Heckman (1998) distinguished the effects of family income from the effects of child and parent ability endowment for five cohorts of Americans, and showed that long term factors are much more important than short-run income affects for the transition to tertiary education. Cameron and Heckman (2001) also estimated a dynamic model of schooling attainment to investigate the sources of racial and ethnic disparity in college attendance. They used National Longitudinal Survey Data (NLSY) data and showed that long term factors account for most of the racial-ethnic college-going differential. In both of their studies, they showed that looking at only a limited or specific time period might increase the importance of credit and borrowing constraints and emphasized the importance of looking at the grade transition model. Recently, Riphahn and Heineck (2009) studied the effect of family background on children's education in Germany. Their findings showed that a strong parental background effect on children's education attainment. They also found that in spite of massive public policy interventions and education reforms to improve "equality of opportunity", there is no significant reduction in the role of parental background for child outcomes over the last decades. Lam et al. (2013) analyzed the large racial gaps in the proportion of high school graduates who enroll in tertiary education in South Africa. They found that controlling for parental background and baseline scholastic ability reduces the estimated impact of household income on university enrollment.

3.3 Ability in Educational Papers

3.3.1 Cognitive Skills

Gintis (1971) defines the meaning of cognitive skills as the individuals capacities to combine, analyze, interpret, and also apply informational symbols. Cawley et al. (2001) mentioned that cognitive ability is a trait partly inherited and partly built through education.

In the U.S., the AFQT (Armed Forces Qualifications Test) and SAT (Scholastic Assessment Test) has been used to proxy cognitive skills. AFQT covers word knowledge, paragraph comprehension, arithmetic reasoning, and mathematics knowledge, and SAT covers critical reading, mathematics and writing. International agencies also evaluates international tests of students' performance in cognitive skills, i.e. TIMSS, PISA, PIRLS and SIMS. Programme for International Student Assessment (PISA) is one of them, which is a worldwide study by the OECD in member and non-member nations of 15-year-old school students' scholastic performance on mathematics, science, and reading.

In a recent study, Hanushek et al. (2008) showed a strong evidence that ignoring differences in cognitive skills significantly changes the relationship between education and economic outcomes. They also emphasized that international comparisons tests on cognitive skills yields much larger skill deficits in developing countries than concluded from just school enrollment data.

3.3.2 Non-Cognitive Personality Traits

Defining and quantifying the non-cognitive personality traits is a challenging process. A change in non-cognitive personality traits has an effect on a change in cognitive skills. Heckman et al. (2006) pointed out that the non-cognitive personality traits are a critical part of human capital, but very hard to measure with precision. While there are several traits taxonomies, the most famous classification of personality traits is the "Big Five" construct of personality, which formulates all personality traits along five uncorrelated dimensions: extraversion, agreeableness, conscientiousness, neuroticism, and openness to experience (Almlund et al., 2011).

- Heckman and Mosso (2014) identify 5 risky behaviors: violent behavior in 1979, tried marijuana before age 15, daily smoking before age 15, regular drinking before age 15, and any intercourse before age 15 as measures of non-cognitive factor.

- Gullone and Moore (2000) study the relationship between personality traits and adolescent risk-behavior. Their research has confirmed four broad groupings of risk-taking behavior, including thrill-seeking, rebellious, reckless and antisocial risk behaviors. Examples include smoking, drinking alcohol, swearing and staying out late.

Other important findings on non-cognitive personality traits are summarized below:

- Cunha and Heckman (2008) proposed a model to represent the evolution of skills as a function of family environments and showed that parental inputs are more influential to determine cognitive skills at early ages and non-cognitive personality traits at later ages. The paper formalized the notion that cognitive skills can promote the formation of non-cognitive personality traits.

- Biglan (2004) showed that the same cluster of adolescents pursued risky behaviors such as antisocial behavior (aggressiveness, violence and criminality), cigarette smoking, alcohol use and the like.

- Heckman et al. (2006) showed that gender interacts with non-cognitive personality traits. For men, non-cognitive traits are valued more in low skill markets and males may benefit more from work friendly non-cognitive personality traits.

Chapter 4

Data and Covariates

In Section 4.1, the specifications of Indonesian Family Life Survey explained. A detailed explanation of how I picked the cohort using IFLS is outlined in Section 4.2. In Section 4.3, important descriptive statistics are summarized. In Section 4.4, I explain the covariates picked for the model.

4.1 Data Set: Indonesian Family Life Survey (IFLS)

Indonesia is the world's fourth most populous country. The nation's capital city is Jakarta. The country shares land borders with Papua New Guinea, East Timor, and Malaysia. Indonesia is home to 300 ethnolinguistic groups (Thomas and Frankenberg, 2001).

The largest ethnic group in Indonesia is the Javanese who make up nearly

Figure 4.1: Indonesia Ethnolinguistics Map

42% of the total population. The Sundanese, Malay, and Madurese are the next largest groups in the country. See Figure 4.1 for Indonesia ethnolinguistics map.

Indonesian Family Life Survey (IFLS) study is a multipurpose survey consisting of demographic, socioeconomic, and health information on individuals, households, and communities. Since Indonesia is a quite heterogeneous country, the design of IFLS is very helpful to capture this heterogeneity. The survey was conducted in 20 different languages.

IFLS has four waves in 1993, 1997, 2000 and 2007. Another wave, which was fielded in 2014-2015, is expected to be available in late 2016. IFLS1 was collected in 1993, which included interviews with 7224 households and with 22,347 individuals (Frankenberg and Karoly, 1995). IFLS2 was conducted in 1997. IFLS2 aimed to reinterview all IFLS1 households and respondents and also to interview all those not interviewed in 1993 (Frankenberg and Thomas, 2000). IFLS3 was conducted in 2000. Over 94% of the households in IFLS1,

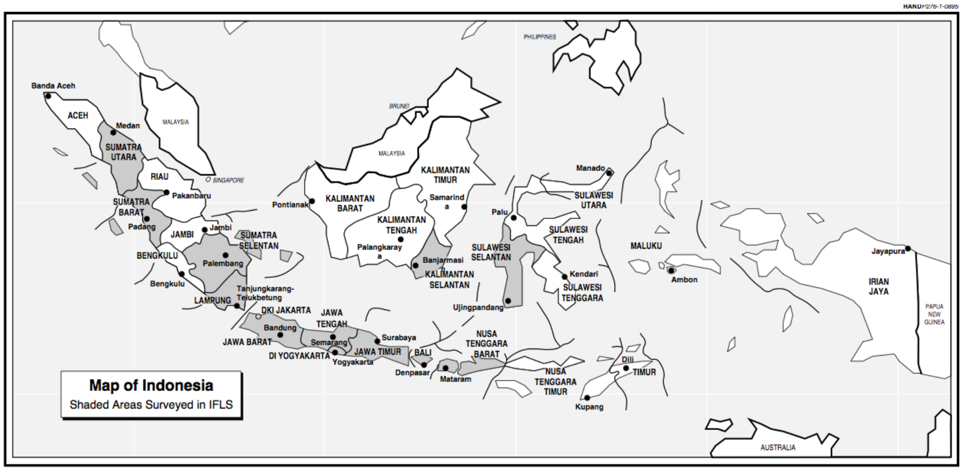


Figure 4.2: Indonesian Family Life Survey

and over 90% of the households in both IFLS1 and IFLS2 were reinterviewed (Strauss et al., 2004). IFLS4 was conducted in 2007. In IFLS4, the recontact rate of original IFLS1 households was 93.6%. A high follow-up rate of IFLS reduces the concerns that can arise from selective attrition.

Over 7,000 households and 30,000 individuals are interviewed in the survey. The sampling scheme for the first wave, which was administered in 1993, is the determinant of the sample in subsequent waves. The IFLS1 sampling scheme stratified on provinces and urban/rural locations, then randomly sampled within strata. Within each stratum, first enumeration areas sampled weighted by population then households are sampled randomly within these enumeration areas. The survey includes 13 of the then existing Indonesia's 27 provinces that contain 83% of population (See Figure 4.2). Within 13 IFLS provinces: four provinces on Sumatra (North Sumatra, West Sumatra, South Sumatra, and Lampung), all five of the Javanese provinces

(DKI Jakarta, West Java, Central Java, DI Yogyakarta, and East Java), and four provinces covering the remaining major island groups (Bali, West Nusa Tenggara, South Kalimantan, and South Sulawesi), 321 enumeration areas were randomly selected based on 1990 Census data (Thomas and Frankenberg, 2001).[1]

The IFLS collected demographic, socioeconomic, migration and education information. Most importantly, the IFLS contains a detailed education information including EBTANAS (Indonesian National Exam) scores after each stage of students' educational paths and Raven-Like test scores in order to measure their cognitive skills. Household survey data are also accompanied by detailed data about the communities from which households are sampled.

In the IFLS, the term *community* refers to a village for rural areas and a neighborhood in an urban setting. The official village leader and a group of his/her staff were interviewed about aspects of community life (Frankenberg and Thomas, 2000). The facility data provides information on the educational services (type of facilities the school has, i.e., library, computer lab, and counseling service) in 312 communities.[2]

The sample for schools is limited to schools which are used by IFLS households. In IFLS survey, schools have students from different communities

[1]For cost-effectiveness and political-violence reasons, 14 of the then existing 27 provinces were excluded.

[2]Nine of which were resided in the same larger community, thus making up 321 communities in total.

not just from one community. Villages might be served by only one well defined school but this is not usually the case for cities. Therefore, the sample frame of the facilities is not constrained to only facilities within the administrative boundary of a village (Amin et al., 2007).

IFLS also provides household analytic weights, which helps to correct for oversampling of urban enumeration areas. When one apply 1993 household weights, the resulting weight reflect the distribution of 13 Indonesian provinces.

4.2 Cohort Selection

I took a sample that consists of 1,776 students who are born between 1978 and 1984 as can be seen in Figure 4.3. This specific cohort is chosen so that when they start school INPRES program (61,000 primary schools were built between 1973 and 1978) and compulsory primary education is fully in effect. My latest data is from IFLS 2007 wave. I took birth year of 1984 latest so that the students have a chance to finish their education fully by 2007.

The cohort sample includes everyone in the IFLS about whom I have information on educational attainment, sex, and year of birth. After eliminations: 1) whose consumption is missing 2) whose location and household id information is missing 3) whose birth order and smoked before age 15 information are missing 4) whose test scores are missing, 1,776 individuals remained in the Cohort Sample. See for details in Table 4.1.

Furthermore, I randomly pick one child from each household to avoid

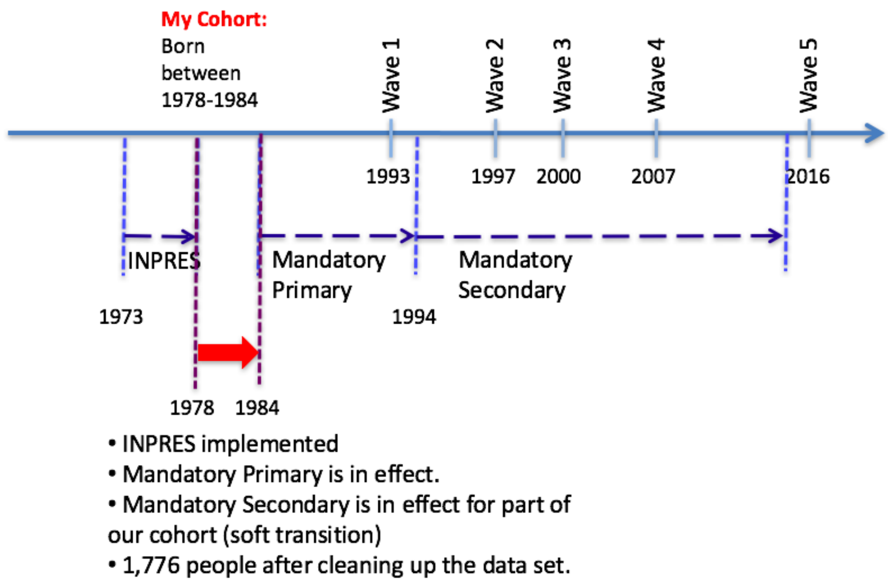


Figure 4.3: Cohort Selection

Table 4.1: IFLS Data Set Construction and Effect of Deletions

Observations	Details
5,545	Core representative of IFLS population
3,468	No consumption
3,410	No province and urban/rural
2,927	No birth order and smoked before 15
2,365	No Raven shape score
1,776	Random sample one child per family

sibling correlations.

My sample consist of individuals who at least finished primary school, i.e. I exclude also individuals who dropped out during primary school, since these students don't have test scores. As a result, I start my model at the end of primary school education.

4.3 Descriptive Statistics

The summary statistics of the final sample is presented in Table 4.2. 79% of the sample continued from primary to lower secondary. If we look at the numbers by gender, 79% of females and 80% of males continued to lower secondary. 50% of females continued to upper secondary and 13% of females attended university, whereas 48% of males continued to upper secondary school and 10% of males attended university, which is slightly less than females for my cohort.

Smoking is gender specific, 22% of males smoke but only a few women smoke, so smoking is dominated by males. 11% of the sample smoked before age 15. Of Indonesian people, 63% of men and 5% of women reported being smokers, a total of 34% of the population (Barber et al., 2008). In Nepal and Indonesia, almost 60 percent of all young males ages 15–24 currently smoke. According to World Development Report young people take more health risks and as they get older the tendency to take risks falls. (WDR, 2007).

As can be seen in Table 4.2, similar number of households are sampled

Table 4.2: Summary Statistics (%)

Variable	Total	Female	Male	Rural	Urban
Low Sec School Attendance	0.79	0.79	0.80	0.72	0.94
Upper Sec School Attendance	0.49	0.50	0.48	0.37	0.75
University Attendance	0.12	0.13	0.10	0.06	0.22
Mother's Education					
No Schooling	0.20	0.20	0.19	0.24	0.12
Some Primary	0.35	0.33	0.37	0.39	0.27
Primary	0.29	0.29	0.29	0.29	0.28
Lower Secondary	0.08	0.09	0.08	0.04	0.16
Upper Secondary	0.05	0.06	0.05	0.02	0.12
University	0.03	0.03	0.02	0.02	0.05
Father's Education					
No Schooling	0.13	0.13	0.13	0.16	0.06
Some Primary	0.30	0.29	0.31	0.35	0.21
Primary	0.29	0.28	0.31	0.31	0.25
Lower Secondary	0.11	0.12	0.11	0.08	0.18
Upper Secondary School	0.12	0.14	0.10	0.07	0.23
University	0.04	0.04	0.04	0.03	0.07
Smoked before 15	0.11	0.00	0.22	0.12	0.07
Obs	1,776	939	837	951	825
Weighted Obs	1,807	937	869	1214	592

from urban and rural areas. After correcting with weights from 1993 survey, the portion of rural increases, which is aligned with the fact that 66% of Indonesia's population lived in rural areas in 1993.[3]

The maximum number of grades for primary school is six years in Indonesia. Children start primary school at the age of 7. Lower secondary and upper secondary school is 3 years. D1 is one year, D2 is two years and D3 is three years of associate degree. University education varies with respect to the field of study. S1 is Bachelor's degree, S2 is Master's degree, and S3 is accepted as doctoral degree. There are religious schools which consists of Islamic schools and Madrasah. There is also adult education in Indonesia for adults to gain new sets of skills and knowledge. There are also kindergartens in Indonesia for ages 4-6 before the primary education. See details in Figure 4.4 (Unesco, 2010).

4.4 Covariates

Covariates are grouped under five categories as can be seen in Table 4.3.

- Personal characteristics: This includes sex and birth order of a child.
- Cognitive abilities and non-cognitive personality traits: See details in Sections 4.4.1 and 4.4.2.
- Family characteristics: I consider log consumption per capita, which is taken from the household consumption data in 1993. I also include

[3] http://data.worldbank.org/indicator/SP.URB.TOTL.IN.ZS?page=4

Table 4.3: Covariates

Variables	Description	Type	Mean (Std)
Personal Characteristics			
Female		Binary	0.52 (0.50)
Birth Order	Order a child is born.	Numerical	2.12 (1.19)
Cognitive Ability			
Shape Tests	Raven-like shape matching questions	Numerical	2.47 (1.20)
Math Score	Ebtanas Test Score after primary school.	Numerical	4.88 (1.70)
Math Score Unknown	Did not remember the result.	Binary	0.35 (0.48)
Non-cognitive Personality Traits			
Ever Smoked	Equal to 1 if a person smoked before age 15.	Binary	0.11 (0.31)
Family Characteristics			
Log Consumption per capita	Household per capita consumption in 1993.	Numerical	13.13 (0.71)
Mother's Education	Highest grade completed, 0 to 17.	Numerical	5.41 (3.79)
Mother's Education Unknown	1 if data is missing.	Binary	0.02 (0.14)
Father's Education	Highest grade completed, 0 to 17.	Numerical	6.28 (4.31)
Father's Education Unknown	1 if data is missing.	Binary	0.06 (0.24)
Location Characteristics			
Rural	Equal to 1 if a student lived in rural area in 1993	Binary	0.67 (0.47)
Java Island	Equal to 1 if a student lived in Java Island in 1993	Binary	0.72 (0.45)
School Characteristics			
Number of Lower Secondary Schools	Number of lower secondary schools used by the community.	Numerical	3.09 (1.80)
Number of Upper Secondary Schools	Number of upper secondary schools used by the community.	Numerical	2.75 (2.84)
Primary School Quality	Residuals of regressing average math score of schools in a community with community's averages of per capita consumption, father education, and mother education levels.	Numerical	0 (0.85)

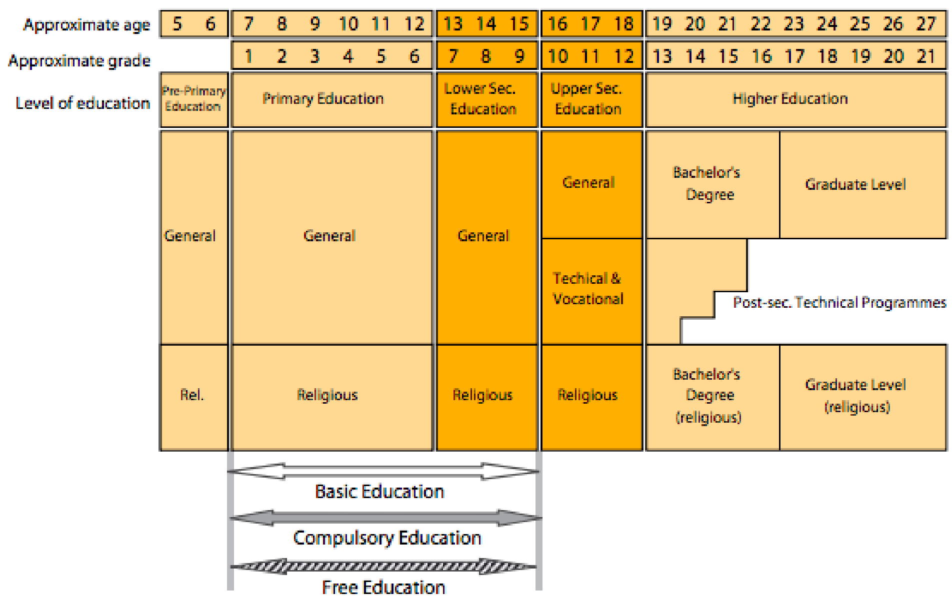


Figure 4.4: Education Structure: Approximate Starting Age and Duration

father's and mother's education as of their highest grade completed. Grade 16 represents the completion of university education. Since a small number of students have this data missing, grades are represented from 1 to 17 and 0 indicates missing data.

- Location characteristics: The job opportunities, wage levels, availability and location of the schools in the area affect the opportunity cost of schooling. I consider two covariates: 1) Whether a person lives in Java Island in 1993 or not. Java island is the most populous island in Indonesia, and most of the prestigious schools were resided in this area. Therefore, I include the effect of living in Java in my covariates. 2) Whether a person lives in a rural area in 1993 or not.

- School characteristics: I consider the number of lower secondary schools and upper secondary schools in each community. The community data contains information on number of lower secondary schools and upper secondary schools used by the community.
- I include birth year dummies to remove their fixed effect.
- I also use covariates to proxy primary school quality. See Section 4.4.3 for more details.

4.4.1 Cognitive Skills

Two main types of cognitive skills affect education preferences, innate math abilities and learned math abilities. In the rest of the section, I will discuss which test scores I use to proxy these abilities, their correlations, and how I address the endogeneity issues associated with test scores.

Cawley et al. (2001) mentioned that cognitive ability is a trait partly inherited and partly built through education. In my analysis I pick two test scores in order to proxy cognitive ability: 1) Raven's matrices shape matching scores to proxy innate math ability, and 2) EBTANAS math score administered at the end of primary school to proxy learned math ability.

Raven's matrices shape matching scores

During IFLS survey, in order to assess the cognitive and math level of individuals, household members between the ages of 7 and 24 were asked to participate in cognitive assessments in IFLS3 and IFLS4 (2000 and 2007 waves). Individuals between the ages 7–14 are asked 12 shape matching questions. Individuals between the ages 15–24 are asked 8 shape matching questions. Shape questions are made up of series of diagrams or designs with a part missing, which are similar to IQ tests called Raven's progressive matrices. Raven Progressive Matrices Test is a classic test of analytic intelligence widely used in both research and clinical settings. Individual differences in the Raven test highly correlate with other complex cognitive tests as well (Jensen, 1987). Among the common IQ tests, Raven's matrices have been shown to have the highest correlation of 0.8 with g-factor (general intelligence factor) (Jensen, 1998). Because of its non-verbal format, it is easier for everyone (young, elderly and patient populations) to take the test. Those taking the tests are expected to select the correct part to complete the designs from a number of options printed beneath (See Figure 4.5). In addition to shape matching questions, math questions are also asked to the individuals taking the survey. I am not using these math questions in my model as these are learned abilities and given they are taken after the schooling decisions are done, they are highly endogenous. Instead I use the shape matching questions.

Given the strong evidence in literature, I pick these shape questions from IFLS administered data as a representative of innate ability. My cohort took this exam in 2000 when they were 16 to 22 years old. Out of 8 questions, I pick 4 questions, EK5, EK6, EK11, and EK12, which are shown in Figure 4.5. I pick these as they are the most discriminative. Number of correct answers for these 4 questions then represent the IQ score.

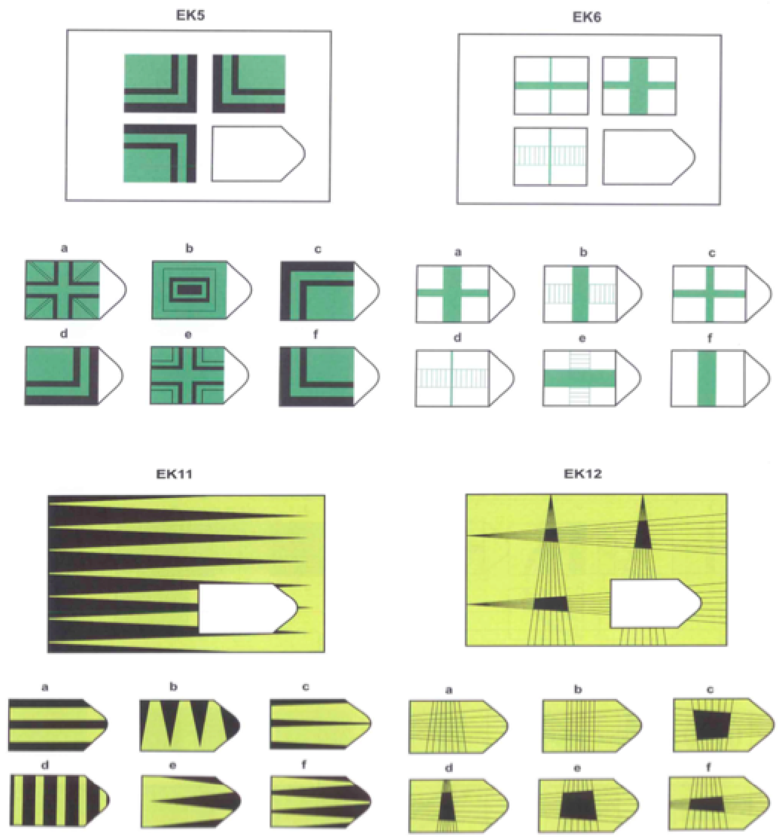


Figure 4.5: Four Questions from the Cognitive Test

Since Raven test is taken at a later age for my cohort, it raises the question if these scores reflect learned abilities or not. Since this is a controversial topic, I will analyze correlations of Raven scores with other tests later in this section.

EBTANAS (National Achievement Test) math scores

For learned ability I pick the nationwide EBTANAS math scores at the end of primary school, which is consistent among individuals as it is taken at the same education level by everyone. National achievement tests (EBTANAS) were administered at the end of each school level, i.e., after primary, lower secondary, and upper secondary. After primary and lower secondary, students are tested on five subjects–PMP (moral and civic education from the nation's five principals), Indonesian, math, science, and social sciences. After upper secondary, students are tested on more detailed subjects depending on what field they pick.[4] Note that EBTANAS was implemented starting in 1980.

All students in my sample set took Ebtanas primary exit test. Only 31% of students indicated the year that they have taken Ebtanas, but did not have their test scores ready during survey. There is no indication in the survey

[4] For example, upper secondary school students majoring in science studies, beside the three main subjects English, Indonesian and Math, they have to take Biology, Physics, and Chemistry. Students majoring in social studies have to take Economics, Sociology, and Geography. While those majoring in language must take History and Anthropology, Indonesian Literature, and one foreign language either French, German, Japanese, Mandarin, or Arabic.

that these missing scores are biased towards a certain group of people. See Table 4.4 for a comparison of Raven Matrices IQ scores of the two group of people. As can be seen in Table 4.4 and Table 4.5, distributions are very similar. To accommodate these students a binary covariate is introduced (math score unknown). In addition, fixed effects of the year test is taken is removed and scores are adjusted accordingly.

Table 4.4: Percentages of IQ scores for missing math scores and available ones

	Raven's Matrices IQ Score				
	0	1	2	3	4
Has Ebtanas score	0.05	0.16	0.23	0.30	0.26
Missing Ebtanas score	0.07	0.16	0.22	0.30	0.25

Table 4.5: Consumption for missing math scores and available ones

	Median Consumption (SEMdn)
Has Ebtanas score	523,138 (22,126)
Missing Ebtanas score	528,437 (39,834)

Correlation of Raven Score and Ebtanas Test Scores

It is important to analyze the correlation of these two test score covariates and justify using two different cognitive scores in my model. As can be seen in Figure 4.6, EBTANAS primary math score and shape questions are loosely correlated, with a Pearson correlation score of 0.15. Correlation with math test in different stages are shown in Table 4.6. Figures 4.6, 4.7, and 4.8 show the detailed analysis of EBTANAS scores (between 0 and 10) to number of shape matching question answered correctly (between 0 and 8). Figure 4.6 shows correlation of Ebtanas primary math and Raven scores, Figure 4.7 lower secondary, and Figure 4.8 upper secondary. Upper left and lower right subfigures show histograms of Ebtanas math and Raven Scores respectively. In all three figures' upper right subfigure, which shows Ebtanas means for each Raven score level, indicates no clear upwards slope that shows more correct shape questions lead to better EBTANAS scores.

Table 4.6: Pearson correlation between EBTANAS math, Raven scores, and IFLS Math Scores

	EBTANAS Math Score After		
	Primary	Lower Secondary	Upper Secondary
Raven Shape Score	0.15	0.11	0.06

As a result, I argue that shape questions even at this late age measure innate ability in line with psychology literature on the connection of Raven scores and analytic innate abilities (Knight and Sabot, 1990) and (Heady, 2003).

I take only Ebtanas primary math scores as covariates and avoid math scores after lower and upper secondary education. Primary math scores are taken before my choice model starts. Moreover as can be seen in Section 5.2, they show no significance for lower secondary school decision when family background controls are present. In light of these findings, it is safe to assume primary math exit scores as exogenous for the choices in my model.

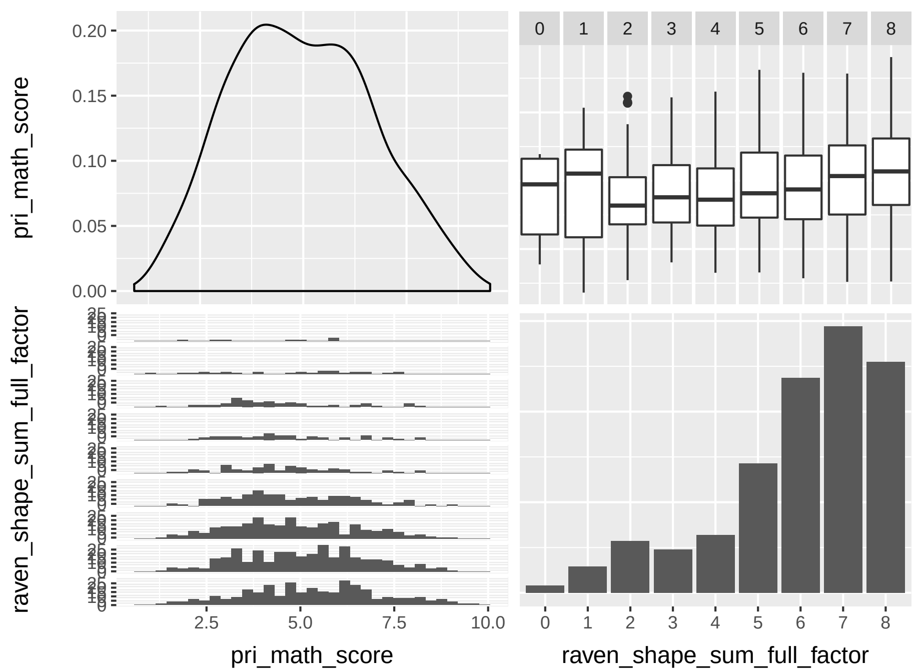


Figure 4.6: Correlation between Primary Ebtanas Math Score and Raven Matrices Score (See text for explanation)

4.4.2 Non-Cognitive Personality Traits

> Intelligence plus character: that is the goal of true education.
>
> *-Martin Luther King Jr.*

Defining and quantifying the non-cognitive personality traits is a chal-

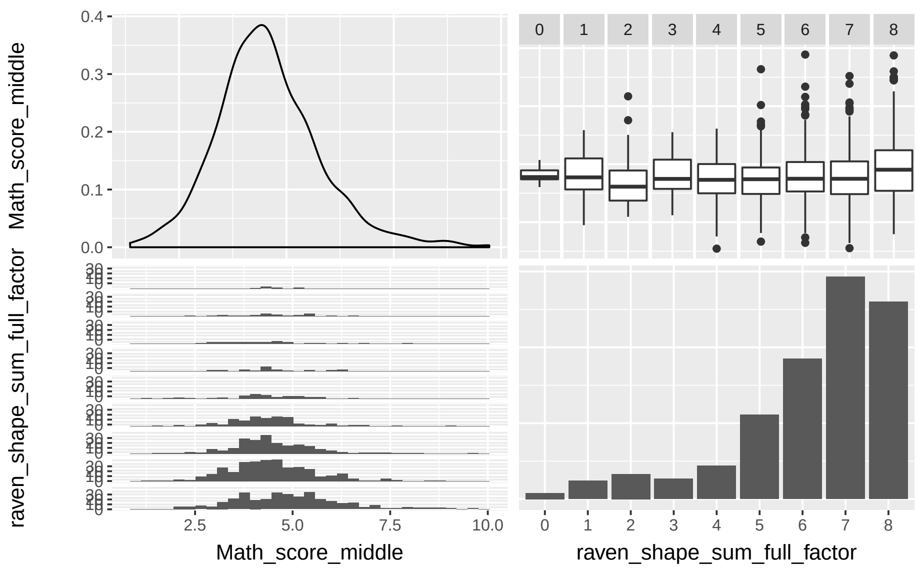


Figure 4.7: Correlation between Ebtanas Lower Secondary Math Score and Raven Matrices Score (See text for explanation)

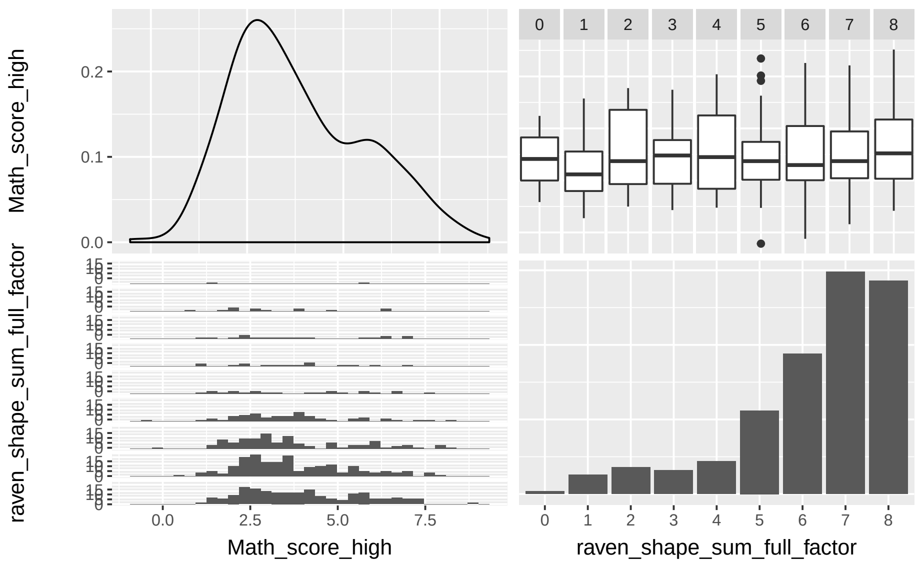


Figure 4.8: Correlation between Ebtanas Upper Secondary Math Score and Raven Matrices Score (See text for explanation)

lenging process. Heckman et al. (2006) pointed out that the non-cognitive personality traits are a critical part of human capital, but very hard to measure with precision. Heckman et al. (2011) identify 5 risky behaviors in their recent research on the United States: violent behavior in 1979, tried marijuana before age 15, daily smoking before age 15, regular drinking before age 15, and any intercourse before age 15 as measures of non-cognitive factor.

As summarized earlier, following the previous literature, I picked smoking before age 15 as a proxy for non-cognitive personality traits. In a 99% Muslim country with highly religious and conservative societal norms, it is difficult to quantitatively track other risky behaviors among those younger than 15 such as regular drinking or intercourse. Note that, according the study of Biglan (2004), the same cluster of adolescents pursued risky behaviors such as antisocial behavior (aggressiveness, violence and criminality), cigarette smoking, alcohol use and the like.

While constructing my covariate, smoking before age 15, I used two questions in the survey. First question is about tobacco habit of the individuals. Second question is about at what age they start to smoke on a regular basis. I only considered the individuals who start smoking before age 15.

Smoking in Indonesia

> A cloud of smoke hovers above his small frame, a cigarette dangling at his lips. As he blows rings high above his head, 14-year-old Faisan explains why he has just bought his third cigarette of the day. "When I have a problem to solve–and I have so many

> problems at school–I have a smoke," he says. "It relaxes me and makes me forget." -*Indonesia's smoking epidemic–an old problem getting younger, The Guardian*

In year 2010, a Youtube video of an Indonesian 2 year old smoking went viral. Child smoking in Indonesia is a big health problem for everyone even if it is more prevalent among the poor.[5] Those who have parents, close relatives or close friends who smoke have a greater chance of smoking (Djutaharta and Surya, 2003).

About one-third of the world's population smokes, mostly in China, India, and Indonesia, in three Asian countries with large populations. With respect to the recent WHO (World Health Organization) report on tobacco use, smoking in Indonesia is among the highest in the world, with 46.8% of males and 3.1% of females aged 10 and over (WHO, 2011). There are strong gender differences in smoking in Indonesia, because female smoking is accepted to be culturally inappropriate.

According to Nawi et al. (2006), smoking also takes place after the circumcision of boys aged 10–12 years in rural areas of Java island, being offered a cigarette during circumcision ceremonies signals a young man's entry into adulthood. The idea that smoking enhances a mans masculinity and is also promoted actively in tobacco advertisements (Nichter et al., 2009). Smoking companies placed billboards with a message for Indonesian kids saying cigarettes are a "cool friend" worth dying for.

[5] http://www.theatlantic.com/international/archive/2015/09/indonesias-marlboro-boys/407308/

Thus, smoking variable is not just a measure of a risky behavior as it is usually accepted in developed countries, but it could be thought of as a combination of cultural influence, smoking advertisement, risky behavior and peer effects, specifically dominant among men living in slums.

4.4.3 School Quality

There is an intensive literature on the school characteristics and quality. As summarized in Section 3.1.2, I first focused on the literature on the school quality and math education. Then, I summarized the recent review done by Glewwe et al. (2011) on the impact of school and teacher characteristics on learning. Following the findings in these studies, I aim to address the importance of primary school quality and its effect on tertiary education in this section.

In order to do that, I use both community and household survey data. The community survey records the math scores on the EBTANAS tests for a random sample of 25 students for each school surveyed.[6] These scores can be used to characterize school's average achievement level.

The school quality a student attended is correlated with the average of 25 random samples of the EBTANAS math scores attained at that school. This average measures how well did the students attending this school on the average done, not a pure measure for school's math education quality. On top of it, there is a selection bias since good students may pick good schools

[6] In primary schools, it was administered with respect to Grade 6, while in lower secondary and upper secondary schools the designated level was Grade 3.

in their community. To remove this school selection bias, I use the average of the math scores of the schools in a community. Given students do not choose the community they live in, we can consider this measure as exogenous for students. There is still a problem that this measure is not a pure measure of school's math education quality. It is correlated with family background, i.e., per capita consumption, father's education, and mother's education of these 25 random students. To remove these effects, I regress average community math score with community average family background controls and pick the residuals to represent community's school quality (See Table 4.7). Given residuals are orthogonal to these background characteristics, they can be taught as average community math score minus average community family background. As can be seen in Table 4.7, per capita consumption and father's education are heavily correlated with community's average math scores. I believe these residuals reflect the quality of primary math education of the students. Table 4.8 also shows that school quality is a strong predictor of students' Ebtanas scores.

I regress average community math score with family background controls as can be seen in Table 4.7:

$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \epsilon$

X_1: Community median log consumption.

X_2: Community average fathers education.

X_3: Community average mothers education.

ϵ: Normal distributed error term.

Table 4.7: Regression on Community's average math scores with community's average family background.

	Estimate	Std. Error	t value	Pr(>|t|)	
Intercept	-7.22	0.82	-8.80	<2e-16	***
Comm. Avg. Log Cons.	1.00	0.07	15.26	<2e-16	***
Comm. Avg. Father's Edu	0.08	0.02	4.02	6e-05	***
Comm. Avg. Mother's Edu	-0.03	0.02	-1.37	0.17	

Signif. codes: 0 *** 0.001 ** 0.01 * 0.05 . 0.1 1

Table 4.8: Regression on EBTANAS primary math scores with school quality, raven score, and family background.

	Estimate	Std. Error	t value	Pr(>|t|)	
Intercept	-1.13	0.92	-1.23	0.22	
Community's School Quality	0.50	0.05	9.33	<2e-16	***
Log Cons.	0.41	0.07	5.67	2e-08	***
Father's Edu	0.02	0.01	1.22	0.22	
Mother's Edu	0.06	0.02	3.76	2e-04	***
Raven Score	0.11	0.04	2.76	0.01	**

Signif. codes: 0 *** 0.001 ** 0.01 * 0.05 . 0.1 1

To summarize, I run three models using above covariates:

- All controls plus number of schools in community.
- All controls except primary math score and raven score plus community average math scores.
    - This is for intuition that primary school may be more important than secondary.
- All controls except primary math score and raven score plus community primary school quality.
    - This is to analyze causal effects of school quality.

Chapter 5

Modelling Indonesia's Educational Attainment

In Section 5.1, sequential choice model is described. In Section 5.2, results are presented.

5.1 Model

In labor economics, sequential discrete choice models are used for different subjects varying from education to retirement. Mare (1980) first showed the importance of sequential choice in the education system and decomposed the final educational attainment into a series of stages in sociology literature. The sequence of grade transition probabilities constructs the probability of schooling attainment and dividing schooling into stages provides in detail analysis in schooling progress. Sequential choice model is also considered by

Cameron and Heckman (1998) and Cameron and Heckman (2001) with an emphasis on dynamic selection bias. Since low ability people drop out more compared to high ability people at early stages of education. Cameron and Heckman (2001) and Cameron and Heckman (1998) used sequential choice model and found that the long-term factors such as parental educational background and child ability to be key factors of continued schooling. They modeled omitted variables such as student ability and motivation as unobserved heterogeneity, which I also follow in this work.

To understand sequential choice model with an example, I use my education model. Students start at primary school seen in Figure 5.1. Students are associated with specific covariates (consumption, urban/rural, etc.), and 4 choices:

- Primary as the highest educational attainment.
- Lower Secondary as the highest educational attainment.
- Upper Secondary as the highest educational attainment.
- Tertiary as the highest educational attainment.

As can be seen in Figure 5.1, there are four levels in our sequential model. The first transition is a choice between primary as the highest grade or lower secondary and above. The second transition is a choice between lower secondary as the highest level or upper secondary and above. The third transition is a choice between upper secondary as the highest or tertiary.

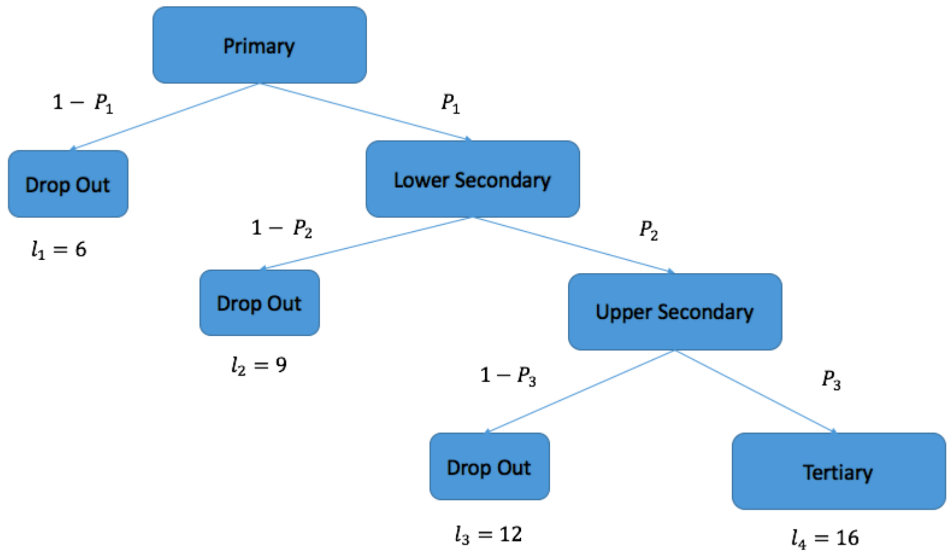


Figure 5.1: Educational Attainment Diagram

5.1.1 Mixture Models Allowing Unobserved Heterogeneity

As we will see in Section 5.2, there is unobserved heterogeneity across students in their distribution in my dataset, even after controlling for the effect of observed variables. Omitted heterogeneity in duration models could lead to misleading inferences about the effects of explanatory variables. To demonstrate the importance of unobserved factors, I start with a model containing no explanatory variables and simple unobserved heterogeneity and follow a hazard model setup.

Suppose a fraction p of the students in our sample have hazard $\lambda_1(t) = \gamma_1$ and fraction $(1 - p)$ have hazard function $\lambda_2(t) = \gamma_2$ where both γ_1 and γ_2 are constants.

With the assumption of population consisting of two types, we sample from a mixture distribution.

$$f(t) = pf_1(t) + (1-p)f_2(t) \tag{5.1}$$

Corresponding hazard function we estimate is not constant in duration t:

$$\lambda(t) = \frac{p\gamma_1 e^{-\gamma_1 t} + (1-p)\gamma_2 e^{-\gamma_2 t}}{pe^{-\gamma_1 t} + (1-p)e^{-\gamma_2 t}} \tag{5.2}$$

It can be shown that $d\lambda(t)/dt < 0$, i.e. negative duration dependence.

As time elapses, the fraction of students from the group with the higher hazard (low ability and motivation) will fall. Because individuals from the other group have a lower hazard (high ability and motivation), the decline in the fraction of individuals from the high-hazard group shows up as a decline in the hazard function over time.

To estimate the model above, one needs to base inference on the mixed distribution resulting from the presence of heterogeneity. Thus, in the example above, the distribution $f(t)$ should be used to form the likelihood function. More generally, the individual densities can be written conditionally on a heterogeneity term v, as $f(t|v)$ and inference can be based on the distribution of observed durations.

$$f(t) = E_v[f(t|v)] = \int f(t|v)p(v)dv \tag{5.3}$$

In any application, the random effect estimator is implemented by assuming a functional form for the structural duration distribution of interest given observed and unobserved variables and a functional form for the distribution of unobservables and also assuming that unobservable is independent of covariates. Maximum likelihood is used to estimate the parameters of the structural duration distribution and the parameters of the distribution of unobservables.

According to Heckman and Singer (1984), provided that there is information about the functional form of a duration model conditional on values of unobserved variables, it is possible to utilize observed duration data to estimate the distribution function of the unobservables and the structural parameters of the model using a nonparametric maximum likelihood procedure.

A flexible approach is to model v as having a discrete distribution, in which case

$$f(t) = \sum_j f(t|v_j)p_j \tag{5.4}$$

where $p_j = Pr(V = v_j)$, $j = 1, ..., J$.

Let the parameter vector $\lambda = (v_1, ..., v_J, p_1, ..., p_J)$. These parameters and J, the number of points in the distribution (called 'mass points'), can be estimated by maximum likelihood. This is the procedure proposed by Heckman and Singer (1984), which shows that any unknown distribution $p(v)$ can be approximated arbitrarily close by such a procedure.

5.1.2 Maximum Likelihood Formulation

My choice model for the individual stages is based on McFadden's random utility. Relative utility of grade progression with respect to dropping out is learned from data using conditional logit like framework.

Using McFadden's random utility model, value of the choice c at stage g is as follows

$$V_g^c = \beta_{g,c} + X_i \beta'_{g,c} + \alpha_{g,c} A_i + \epsilon_{g,c} \tag{5.5}$$

where X_i is observed variables, $\beta_{g,c}$ is the intercept, $\beta'_{g,c}$ is vector of coefficients (note that to identify the model, coefficients of one of the choices will be set to 0), A_i is the ability, and $\epsilon_{g,c}$ is an extreme value distributed error term.

The theory states that a person decides whether to continue or drop out of school after evaluating a marginal benefit and marginal cost calculation. The individual will choose to continue in school whenever the net benefit of doing so is positive, i.e. $V_g^{prog} > V_g^{drop}$, stage is progressed, otherwise drop out. To identify the model coefficients of V_g^{drop} are set to zero, i.e. $V_g^{drop} = \epsilon_{g,drop}$. Since only one of the choices is left with coefficients associated to it, I will drop the index c from the equations in the following text. The probabilities that person i drops out or progresses can be written as

$$P(drop|X, A) = \frac{1}{1 + e^{\beta_g + X\beta'_g + \alpha_g A}} \tag{5.6}$$

$$P(prog|X, A) = 1 - P(drop|X, A) \tag{5.7}$$

These probabilities define a schooling transition model.

Likelihood without Unobserved Heterogeneity

If we assume that there is no unobserved heterogeneity that connects each educational level to each other, maximum likelihood without unobserved heterogeneity is the summation of log likelihoods of each individual at each stage.

$$\arg\max_{\alpha,\beta} ln(L_g) = \sum_i \sum_g \sum_c ln(P(c|X_i, A_i)^{d_g^c}) \tag{5.8}$$

$$d_g^c = \begin{cases} 0 & \text{if c is not chosen} \\ 1 & \text{if c is chosen} \end{cases}$$

Likelihood with Unobserved Ability

As indicated earlier, omitted variables such as "ability" and "motivation" may cause dynamic selection bias. Perceived ability can change over time in early schooling as child develops and higher ability people may find school less difficult and be more likely to progress. As grades increase, low ability students drop out more than high ability students, changing the mixture of the distribution and causing dynamic selection bias.

As a result, while analyzing policy effects, one should also need to consider the unobservable variables. There are different ways to deal with unobservable data. If individuals have been randomly assigned to the treatment and control groups then this information could be used to control for unobservables. Another way is to use instrumental variables.

I follow the unobserved heterogeneity approach described earlier. This model considers omitted variables such as "ability" and "motivation", which may cause dynamic selection bias. Ability connects decision process in all stages. Since it is not observed, we need to integrate out A using its distribution.

An individuals likelihood can be written as:

$$L_i = \int_A p(A) \prod_g \prod_c P(c|X_i, A)^{d_g^c} dA \tag{5.9}$$

We can model $p(A)$ as a discrete distribution with a few mass points (Heckman and Singer, 1984). In this case, each population type is assigned a percentage.

Maximum likelihood yields estimates for the average values of the parameters β and α. If there are several distinct data-generating processes, estimating a single set of parameters is inappropriate and may lead to wrong results. Because of that, I estimate using a discrete distribution with mass points in order to account for heterogeneity as described above. The estimation procedure yields estimates of the relative sizes of the different groups, as well as the group-specific parameters.

Then maximum likelihood of grade progression with **unobserved ability** can be written as:

$$\arg\max_{\beta,\alpha,\pi} ln(L) = \sum_i ln \sum_A P(A = a) \prod_g \prod_c P(c|X_i, A = a)^{d_g^c} \tag{5.10}$$

where $P(c|X_g, A)$ is given in Equation 5.6 and 5.11 and A is the ability types e.g. for two mass points these could be thought as low ability and high ability respectively. Note that additive separability of log is lost now. $p(A)$ is approximated by a discrete distribution with mass points (with respect to finite mixture model, number of mass points is fixed and bounded). If $P(A = a)$ is the probability of someone belonging to that type, then $\sum_A P(A = a) = 1$. Note that $P(A = a)$ does not depend on covariates. In literature, usually

2 or 3 types is found to be adequate for most data. I also observe similar results with my dataset. The value of progressing stage g for ability type A is as follows:

$$V_{g,A}^{prog} = \beta_g + X_i\beta_g^{'} + \alpha_g A + \epsilon_{g,A}^{prog} \tag{5.11}$$

Since in the utility calculation, the component $\alpha_g A$ is a constant given ability type, we can further simplify the model by combining this with the model intercept. This will result one intercept per ability type. It means to estimate N_A intercepts per stage. Equation 5.11 can be rewritten to represent the utility of type A as:

$$V_{g,A}^{prog} = \beta_g^A + X_i\beta_g^{'} + \epsilon_{g,A}^{prog} \tag{5.12}$$

To summarize, ability specific utilities vary in their intercepts only, which corresponds to assuming that the importance of observed covariates are same for all ability types. Also $\beta_g^{'}$ is stage specific, i.e. importance of covariates are different for each stage.

By moving from single type maximum likelihood problem (Equation 5.8) to multi-type problem (Equation 5.10), nice properties of a concave function is lost. Directly maximizing the full likelihood is also challenging due to the lack of additive separability and requirement of multiplication of small probabilities, which could introduce instabilities, slow convergence as well as convergence to local maxima.

Unobserved Ability Distribution for Population Subsets

I assumed that ability type distribution is same for all population while explaining the likelihood with unobserved ability. It is possible that sub-populations might have different ability distributions.

Let's consider a discrete covariate X_k (e.g. Raven shape scores) and different ability distributions for its values, then the equation can be written as follows:

$$\arg\max_{\beta,\alpha,\pi} ln(L_g) = \sum_i ln \sum_A P(A = a|X_{i,k} = x_k) \prod_g \prod_c P(c|X_{i,\neq k}, A = a)^{d_g^c} \tag{5.13}$$

where $X_{i,\neq k}$ is all X excluding X_k.

Note that X_k is removed from likelihood for better identification of the unobserved ability. By calculating the ability distribution for population subsets, we may increase the power of unobserved ability part.

Forward Looking Agents

Heckman et al. (2016) asks the question if agents are forward-looking or not. Their analysis shows that a high school student does not consider benefits of attending college when deciding whether to graduate from high school or drop out earlier. On the other hand, they observe that decision to enroll in college or not considers the value of graduating from college. They conclude that high school students trying to decide to graduate or not either are not forward looking or their abilities to process publicly available information is weak. They then mention that this implicit assumption in Bellman-equation-based education models call their conclusions into question.

My model is not at the granularity that I separate enrolling and graduation as two choice stages. This is done to avoid complicating the model and increase its sensitivity since most people who enroll also graduate in my dataset. Following conclusions in Heckman et al. (2016), at the choice granularity of my model, my assumption that agents are not forward looking is reasonable.

Earnings and Future Value at Absorbing States

In my model, I avoid using future expectations on income and wages and let the relative utility of choices handle this (Equation 5.5). There are several reasons for this:

- Different from many other studies, instead of earnings, my outcome variable of interest in this dissertation is tertiary educational attainment, which is very low in Indonesia.
- Future value from wages may not be a good metric in a developing country. Self employment is common and many women don't work but get education to find a good husband.
- My model starts at very early stages of education. At these early stages, parents are heavily involved in education decisions and it is not clear what objective for future expectations is optimized.

5.1.3 Maximum Likelihood Estimation

There are three possible options for solving Equation 5.10: a) constrained optimization, b) unconstrained optimization, and c) expectation maximization (EM). Constrained optimization solves Equation 5.10 directly along with constraints of π_A. With some variable transformation tricks, maximization can be converted into an unconstrained optimization. In both of these cases additive separability of log and concavity of the function maximized is lost. EM on the other hand is able to transform the problem into a series of additively separable and concave maximization problems.

Constrained Optimization

Equation 5.10 is a constrained optimization with constraints such that $0 \leq P(A = a) \leq 1$ and $\sum_A P(A = a) = 1$. Optimization packages with box constraints could be used to maximize this likelihood, for example Fortran PORT library (nlminb in R). [1] In my attempts, solving constrained optimization did not return good results most of the time, i.e., usually a local maximum is returned. A better approach is to reformulate Equation 5.10 as an unconstrained optimization.

Unconstrained Optimization

Instead of constrained maximization, we could use the following approach that enforces both constraints implicitly. The derivative of $\arg\max ln(L_g)$

[1] https://stat.ethz.ch/R-manual/R-devel/library/stats/html/nlminb.html

with respect to the mixing probabilities $P(A = a)$ is required in constrained maximization because the values of $P(A = a)$ are constrained to being probabilities and adding up to one. This constraint can be handled by writing variables $P(A = a)$ in turn as functions of unconstrained variables γ_a as follows:

$$P(A = a) = \frac{e^{\gamma_a}}{\sum_{a=1}^{A} e^{\gamma_a}} \tag{5.14}$$

Derivatives can be calculated easily as well.

$$\frac{\partial P(A = a)}{\partial \gamma_j} = \begin{cases} P(A = a) - P(A = a)^2 & \text{if a = j} \\ -P(A = a)_j P(A = a)_a & \text{otherwise} \end{cases} \tag{5.15}$$

Among the optimization packages I tried, R's *nlm* function, which uses a Newton-type algorithm, seems to be the most robust among them. *nlm* also has a nice feature; if you provide an explicit gradient function, it does a numerical check to verify its correctness and warns the user. While more robust, unconstrained optimization converges to local maxima many times as well, so multiple re-runs with random initial values and picking the maximum likelihood among them is necessary. As initial values I use zero mean normal random values for coefficients and equal probabilities for $P(A = a)$.

While this approach addresses the issues with constrained optimization, it does not address the loss of additive separability and the need to multiply small probabilities in log likelihood function.

EM Algorithm

Expectation Maximization (EM) is a tool to estimate the parameters of a finite mixture model from a set of data points. It is specifically designed for the case of unobserved types, which is what I try to solve. EM can be used to maximize the log likelihood in Equation 5.10 without loosing additive separability of log. EM and Newton's methods, which is used in unconstrained optimization, have similarities in their mechanism.[2] To understand the difference between EM and Newton's method, let me compare the maximization of a scalar function $f(\theta)$ of a vector θ with both of them.

- Newton Method: Starting at a given point $\theta(0)$, approximates f in a neighborhood of $\theta^{(0)}$ with a paraboloid $g_0(\theta)$ and then finds the maximum of the paraboloid by solving a system of linear equations to obtain a new point $\theta^{(1)}$. This approximation continues for $\theta = \theta^{(1)}, \theta^{(2)}, ...$ until the change from $\theta^{(i-1)}$ to $\theta^{(i)}$ is small enough. This means the convergence to a point θ^*, a local maximum of f.

- EM Method: EM finds a new function $g_i(\theta)$, which is a lower bound for f such that $g_i(\theta) \leqslant f(\theta)$ in the neighborhood of the current estimate $\theta^{(i)}$. g_i is generated using Jensen's inequality, which converts log of sums into sum of logs, and as a result, g_i is an additively separable concave function. The function g_i and f touch at $\theta^{(i)}$, i.e. $f(\theta^{(i)}) = g_i(\theta^{(i)})$. In EM case, g_i need not be a paraboloid (In Newton's method paraboloid is not necessarily a lower bound of f). $\theta^{(i+1)}$ is then picked as the maximum of

[2] http://www.cse.psu.edu/ rtc12/CSE586/papers/emTomasiTutorial.pdf

g_i. $f(\theta^{(i+1)}) \geqslant g_i(\theta^{(i+1)}) \geqslant g_i(\theta^{(i)}) = f(\theta^{(i)})$ so that $f(\theta^{(i+1)}) \geqslant f(\theta^{(i)})$ (Figure 5.2). Instead of one complicated maximization step, EM maximizes a simpler problem, g_i, which is additively separable, several times by iteratively maximizing the likelihood. Each iteration is guaranteed to increase the likelihood (Dempster et al., 1977) (See Figure 5.2). Recent econometric work has also been utilizing EM for solving models with unobserved types[3].

For both methods, convergence is guaranteed, but there is not a known fact for which method will give a better result. It depends on the function f and the lower bound g_i. If one finds bound functions g_i that are very similar to f, then it is possible that EM works better than Newton's method.

Dempster et al. (1977) showed that the likelihood never decreases from one EM iteration to the next iteration. EM algorithm is a hill climbing algorithm. Therefore, it converges monotonically to the maximum. EM algorithm iterates on two steps.

At iteration i:

- Expectation (E) step calculates estimates for unobserved types. The conditional probability of each observation being in each unobserved state is calculated given the data and the coefficients of the model from previous iteration.

- Maximization (M) step formulates and maximizes g_i by taking ability type assignments from expectation step as observed.

[3] Arcidiacono (2005), A Bruhin (2010).

See Appendix A for the derivation of EM formulas.

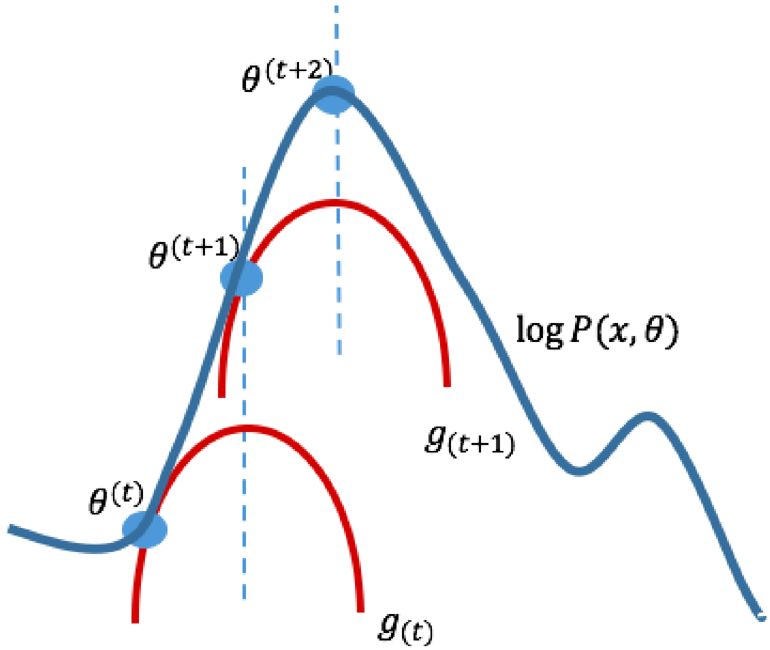


Figure 5.2: Convergence of the EM Algorithm

5.2 Results

5.2.1 Estimation Strategy

To estimate the coefficients of Maximum Likelihood (Equation 5.10), I follow the following strategy. First EM is run to get close to the maximum. At each EM iteration amount of increase in likelihood is checked and if the increase is below a threshold, then EM stops. At that point, I feed coefficients estimated by EM to unconstrained optimization as initial values to find the final maximum. This still might be a local maximum. I run maximum likelihood estimation multiple times with different random initial values sampled from a zero mean normal distribution to make sure I found the global maximum.

Optimization packages are able to calculate numerical estimates of gradients if non provided. Unfortunately, this significantly increases run time and makes it infeasible to rerun maximum likelihood estimation many many times with different initial values. As a result, I chose to provide analytic gradients, which significantly reduces the run time. Single maximum likelihood estimation takes one minute on my laptop. Then standard errors are generated using the Hessian of the full likelihood (Equation 5.10) around the coefficient estimates corresponding to maximum likelihood. This takes about two minutes on my laptop.

5.2.2 Maximum Likelihood Estimates

Solving Equation 5.8 results the estimates for the model without unobserved heterogeneity (one type assumption) and solving Equation 5.10 results the estimates for the model with unobserved heterogeneity (multiple types assumption). Table 5.1 shows the likelihoods estimates for one, two, and three ability type assumption as well as the p-value of their likelihood ratio tests. From these numbers, a two type model with unobserved heterogeneity seems to be the best fit for my dataset. As can be seen, the likelihood ratio test corresponding to going from two to three types shows very high p-value indicating that two mass points are enough to approximate the distribution of unobserved ability.

Table 5.1: Maximum likelihoods and p values for different number of types

Number of Types	Maximum Likelihood	π_A	p-value of LR Test
No Unobservables	-1721.48	1	-
Two Types	-1715.08	0.40, 0.60	0.01
Three Types	-1712.71	0.08, 0.37, 0.55	0.32

Model Coefficients with and without Unobserved Ability

Maximum likelihood estimates of Equations 5.8 and 5.10 are presented in Tables 5.2, 5.3, and 5.4 each corresponding to the three stages of my model. As can be seen in these tables, in the first two stages of the model, coefficients of some of the covariates are highly accentuated in the two type model. For example, the increase of the odds ratios from 1-type to 2-type model coefficients for lower secondary to upper secondary transition are as follows: -living in a rural area at age 12: 43% lower odds ratio, -smoking before age 15: 42% lower odds ratio, and -log consumption per capita: 32% higher odds ratio. On the other hand, coefficients stayed very close in the third stage. This demonstrates the correction for selection bias. At the third stage, most of the low ability students dropped out earlier leaving only high ability students, which explains the similarity of 1-type and 2-type models at this stage. This can also be observed by looking at the type specific intercepts. In the first two stages the separation of intercepts is higher than it is in the third stage.

Table 5.2: From Primary to Lower Secondary School

	One Type Without Unobserved Heterogeneity Coefficient (Standard Error)	Two Types With Unobserved Heterogeneity Coefficient (Standard Error)
Personal Characteristics		
Sex	** -0.33 (0.17)	* -0.37 (0.21)
Birth Order	0.08 (0.07)	0.14 (0.09)
Cognitive Ability		
Raven Test Scores	*** 0.36 (0.06)	*** 0.47 (0.10)
Math Score	0.06 (0.06)	0.07 (0.08)
Math Score Unknown	0.24 (0.36)	0.23 (0.45)
Non-Cognitive Personality Traits		
Ever Smoked	** -0.51 (0.25)	* -0.60 (0.33)
Family Characteristics		
Log consumption per capita	*** 0.62 (0.13)	*** 0.77 (0.18)
Mothers Education	*** 0.12 (0.03)	*** 0.16 (0.05)
Mothers Education Unknown	** 1.41 (0.65)	** 1.91 (0.80)
Fathers Education	*** 0.18 (0.03)	*** 0.22 (0.04)
Fathers Education Unknown	* 0.49 (0.30)	0.58 (0.40)
Location Characteristics		
Rural	*** -0.78 (0.18)	*** -0.93 (0.24)
Java Island	** -0.45 (0.16)	** -0.46 (0.20)
School Characteristics		
Number of Lower Secondary Schools	** 0.10 (0.05)	** 0.12 (0.06)
Number of Upper Secondary Schools	–	–
Type 1	*** -9.06 (1.70)	*** -9.31 (2.46)
Type 2		*** -12.56 (2.70)
π_A		*** 0.40 (0.12)
*** 0.01 ** 0.05 * 0.1		

Table 5.3: From Lower Secondary School to Upper Secondary School

	One Type Without Unobserved Heterogeneity Coefficient (Standard Error)	Two Types With Unobserved Heterogeneity Coefficient (Standard Error)
Personal Characteristics		
Sex	-0.12 (0.14)	-0.21 (0.19)
Birth Order	0.08 (0.05)	* 0.13 (0.08)
Cognitive Ability		
Raven Test Scores	*** 0.19 (0.06)	*** 0.34 (0.09)
Math Score	*** 0.15 (0.05)	*** 0.20 (0.07)
Math Score Unknown	* 0.51 (0.30)	* 0.69 (0.42)
Non-Cognitive Personality Traits		
Ever Smoked	*** -0.77 (0.23)	*** -1.12 (0.35)
Family Characteristics		
Log consumption per capita	*** 0.54 (0.11)	*** 0.82 (0.17)
Mothers Education	*** 0.12 (0.03)	*** 0.18 (0.04)
Mothers Education Unknown	*** 1.20 (0.45)	*** 1.79 (0.63)
Fathers Education	*** 0.15 (0.02)	*** 0.22 (0.04)
Fathers Education Unknown	*** 1.21 (0.32)	*** 1.64 (0.48)
Location Characteristics		
Rural	*** -0.70 (0.14)	*** -1.06 (0.23)
Java Island	*** -0.43 (0.14)	*** -0.62 (0.20)
School Characteristics		
Number of Lower Secondary Schools	–	–
Number of Upper Secondary Schools	** 0.07 (0.03)	** 0.11 (0.05)
Type 1	*** -8.77 (1.46)	*** -12.13 (2.41)
Type 2	–	*** -15.03 (2.83)
π_A		*** 0.40 (0.12)
*** 0.01 ** 0.05 * 0.1		

Table 5.4: From Upper Secondary School to University

	One Type Without Unobserved Heterogeneity Coefficient (Standard Error)	Two Types With Unobserved Heterogeneity Coefficient (Standard Error)
Personal Characteristics		
Sex	** 0.40 (0.18)	** 0.41 (0.19)
Birth Order	-0.10 (0.07)	-0.09 (0.07)
Cognitive Ability		
Raven Test Scores	** 0.22 (0.09)	** 0.25 (0.10)
Math Score	*** 0.37 (0.07)	*** 0.40 (0.08)
Math Score Unknown	*** 2.75 (0.48)	*** 2.88 (0.56)
Non-Cognitive Personality Traits		
Ever Smoked	** -0.98 (0.46)	** -1.07 (0.49)
Family Characteristics		
Log consumption per capita	*** 0.49 (0.14)	*** 0.56 (0.17)
Mothers Education	*** 0.14 (0.03)	*** 0.15 (0.03)
Mothers Education Unknown	0.26 (0.69)	0.36 (0.72)
Fathers Education	*** 0.09 (0.03)	*** 0.11 (0.04)
Fathers Education Unknown	- 0.20 (0.55)	-0.10 (0.58)
Location Characteristics		
Rural	-0.04 (0.20)	-0.11 (0.22)
Java Island	-0.10 (0.18)	-0.13 (0.20)
School Characteristics		
Number of Lower Secondary Schools	–	–
Number of Upper Secondary Schools	–	–
Type 1	*** -12.25 (1.93)	*** -13.15 (2.44)
Type 2	–	*** -14.18 (3.17)
π_A		*** 0.40 (0.12)
*** 0.01 ** 0.05 * 0.1		

Analyzing the model coefficients, it can be seen that parental background, region of residence, number of schools, and child's cognitive and non-cognitive personality traits are key factors of educational path decisions. Household per capita consumption, father's education, and mother's education are important in all stages of educational choices. Living in a rural areas as opposed to an urban area negatively affects the earlier stages of educational path but does not seem to have an effect last stage. Being female has a negative effect on lower secondary school decision, and has a positive effect on tertiary decision. Birth order has a positive effect for second stage. Innate ability affects all stages of the educational path. Non-cognitive personality trait, i.e. smoking before age 15, has a negative effect for all stages.

Low Ability vs. High Ability Individuals

Two type model splits the population into two types, first corresponding to 40% and the second 60%. These two groups (types) behave very differently. Table 5.5 and Figure 5.3 shows educational attainments of these two groups. I call the first group as high ability type and the second group low ability type. As can be seen, even for high ability individuals my model predicts fairly low probability of attending tertiary education level.

Model Coefficients with School Quality Covariate

A striking result we observed in the coefficients at Table 5.4 is that primary math scores have a direct effect on whether to continue to tertiary education or not. Usually most studies focus on the indirect effect of early educa-

Table 5.5: Predicted Education Attainment Percentages

	Primary	Lower Secondary	Upper Secondary	Tertiary
High Ability Type	0.04	0.21	0.58	0.18
Low Ability Type	0.30	0.36	0.27	0.08

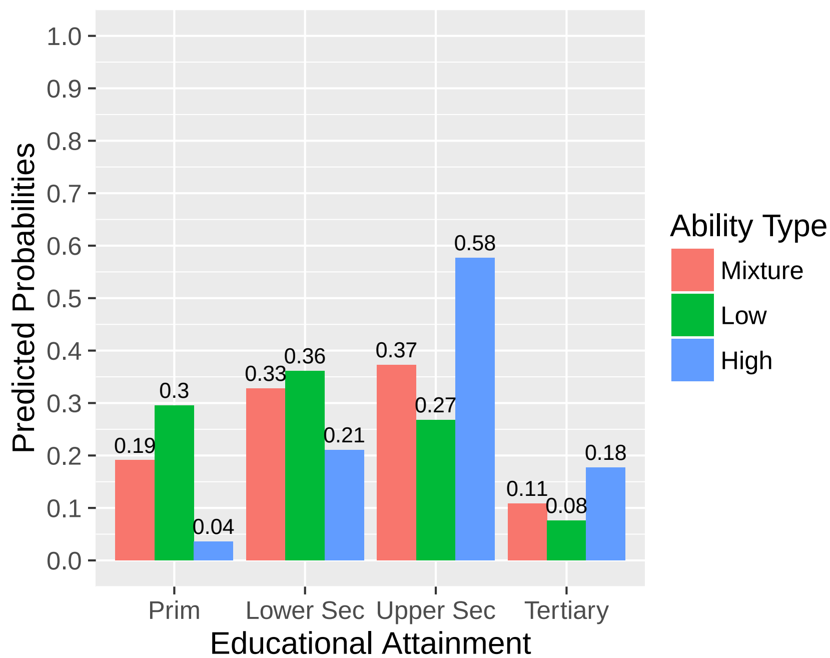


Figure 5.3: Predicted probabilities for educational attainment for high and low ability types.

tion, i.e. better primary education will lead to more students continuing to lower secondary, which in turn will lead higher tertiary enrollments. Having observed a direct effect brings the question if fundamental math abilities acquired in early education have a big impact on whether the student goes to university. This motivates the following analysis in this section regarding a causal relationship between early math education quality and bringing up highly educated individuals.

Primary math test scores are correlated with student's abilities as well as school's math education quality. An interesting question is, if we isolate math education quality, would we still observe same direct impact on the tertiary transition. For each school visited in the survey, IFLS provides math test scores of randomly picked 25 students. These scores can be averaged to create an indicator of how well the students attending that school on the average do. There is one selection bias to consider before we can use this as a school quality metric: good students probably pick the better school among their nearby alternatives.

To remove this primary school selection bias, we can average school's average math score over the schools in a community, which creates a community level average math score. This score now is exogenous for a student's educational choices since the student does not choose the community he/she lives in. On the other hand, we still need to check if the community math average is still a significant covariate for tertiary transition. Table 5.6 shows the coefficients after removing students math scores but adding instead community average math scores for primary, lower secondary, and upper secondary.

As can be seen, even after averaging out over all schools in the community, primary math score is significant for the tertiary transition.

As discussed in Section 4.4.3 in detail, community average math scores are correlated with family background and as a solution I created a new covariate I call community school math quality by regressing with community's average family background and keeping the residuals. Table 5.7 shows that community's school quality still has a direct effect on tertiary transitions even after controlling for family background correlations.

These findings further support the importance of promoting cognitive ability and high quality education early in life. The results found so far suggest that long-term factors may have a stronger influence on the determinants of schooling attainment. Previous literature, Cameron and Heckman (1998) also pointed out the importance of long term factors as opposed to short term credit constraints.

Table 5.6: Educational Path Transitions and Average Community School Math Scores

	Lower Secondary	Upper Secondary	University
Personal Characteristics			
Sex	** -0.81 (0.35)	-0.19 (0.15)	** 0.43 (0.18)
Birth Order	0.22 (0.15)	0.08 (0.06)	* -0.13 (0.07)
Non-Cognitive Personality Traits			
Ever Smoked	** -1.11 (0.51)	*** -0.92 (0.27)	** -1.05 (0.46)
Family Characteristics			
Log consumption per capita	***1.09 (0.28)	*** 0.60 (0.14)	*** 0.57 (0.14)
Mothers Education	*** 0.32 (0.09)	*** 0.14 (0.04)	*** 0.17 (0.03)
Mothers Education Unknown	** 3.13 (1.31)	*** 1.51 (0.51)	0.71 (0.70)
Fathers Education	*** 0.32 (0.08)	*** 0.16 (0.03)	*** 0.09 (0.03)
Fathers Education Unknown	0.58 (0.59)	*** 1.16 (0.33)	-0.22 (0.55)
Location Characteristics			
Rural	*** -1.34 (0.46)	*** -0.72 (0.19)	0.08 (0.20)
Java Island	*** -0.93 (0.35)	*** -0.48 (0.16)	-0.07 (0.19)
Average Com. Pri. Math Score	*** 0.46 (0.17)	** 0.17 (0.08)	*** 0.32 (0.11)
Average Com. Low Sec. Math Score	–	0.08 (0.07)	0.09 (0.08)
Average Com. Upper Sec. Math Score	–	–	-0.002 (0.04)
School Characteristics			
Number of Lower Secondary Schools	** 0.18 (0.08)	–	–
Number of Upper Secondary Schools	–	** 0.08 (0.04)	–
Type 1	*** -20.87 (4.53)	*** -10.95 (2.77)	*** -13.71 (2.52)
Type 2	*** -15.75 (3.81)	*** -9.93 (1.96)	*** -13.06 (2.10)
π_A	*** 0.30 (0.06)	*** 0.30 (0.06)	*** 0.30 (0.06)
*** 0.01 ** 0.05 * 0.1			

Table 5.7: Educational Path Transitions and School Quality

	Lower Secondary	Upper Secondary	University
Personal Characteristics			
Sex	** -0.82 (0.35)	-0.18 (0.16)	** 0.43 (0.18)
Birth Order	0.22 (0.15)	0.08 (0.06)	* -0.12 (0.07)
Non-Cognitive Personality Traits			
Ever Smoked	** -1.16 (0.51)	*** -0.93 (0.26)	** -1.03 (0.45)
Family Characteristics			
Log consumption per capita	*** 1.20 (0.28)	*** 0.63 (0.14)	*** 0.63 (0.14)
Mothers Education	*** 0.33 (0.10)	*** 0.14 (0.03)	*** 0.17 (0.03)
Mothers Education Unknown	** 3.27 (1.35)	*** 1.48 (0.51)	0.67 (0.69)
Fathers Education	*** 0.34 (0.08)	*** 0.16 (0.03)	*** 0.09 (0.03)
Fathers Education Unknown	0.65 (0.59)	*** 1.17 (0.32)	-0.20 (0.55)
Location Characteristics			
Rural	*** -1.56 (0.48)	*** -0.82 (0.19)	-0.09 (0.19)
Java Island	** -0.91 (0.35)	*** -0.46 (0.15)	-0.06 (0.19)
School Characteristics			
Community's School Math Quality	* 0.34 (0.18)	0.08 (0.08)	** 0.27 (0.11)
Number of Lower Secondary Schools	** 0.19 (0.08)	–	–
Number of Upper Secondary Schools	–	** 0.08 (0.04)	–
Type 1	*** -19.65 (4.29)	*** -9.90 (2.57)	*** -12.11 (2.36)
Type 2	*** -14.43 (3.64)	*** -8.94 (1.80)	*** -11.45 (1.95)
π_A	*** 0.29 (0.06)	*** 0.29 (0.06)	*** 0.29 (0.06)
*** 0.01 ** 0.05 * 0.1			

Chapter 6

Policy Simulations

6.1 Policy Simulations

My objective in this dissertation is to analyze what affects the educational attainment, with a specific focus on tertiary attainment in a developing country. At Section 5.2, all important factors were discussed. Now in the simulation section, using these important covariates, I try to understand: a) the effect of consumption increase on each educational stage, b) the effect of increasing the number of middle and high schools on tertiary education with respect to primary math score quartiles, consumption quartiles, and urban or rural location, c) equating the effect of school quality in Java Island to other islands, d) the effect of school quality in different locations in Indonesia.

The main result of this chapter is that improving school quality is much more significant for my cohort than building more schools. Due to Indonesia's push for building schools near the students over the years, for my cohort born

1978–1984, the need for better education became more immediate.

Building school experiment is done by adding one more school to each community. In the current system, schools have been built with respect to the need in each community. Thus, I try to understand how adding one more school in each community–a similar average capacity school of this community–affect the educational attainment. The ideal experiment would be to do a random experiment and build a school to some communities and not to others, and see the effect of adding one more school by comparing the educational attainment of these communities. Almost all communities have middle and high schools built when my cohort were school age. As a result this experiment mostly measures relaxing the crowding in schools. As future work, more analysis on school construction could be done with a focus on class size, teacher per class ratio and on reducing the multiple shifts with a more detailed data on these characteristics.

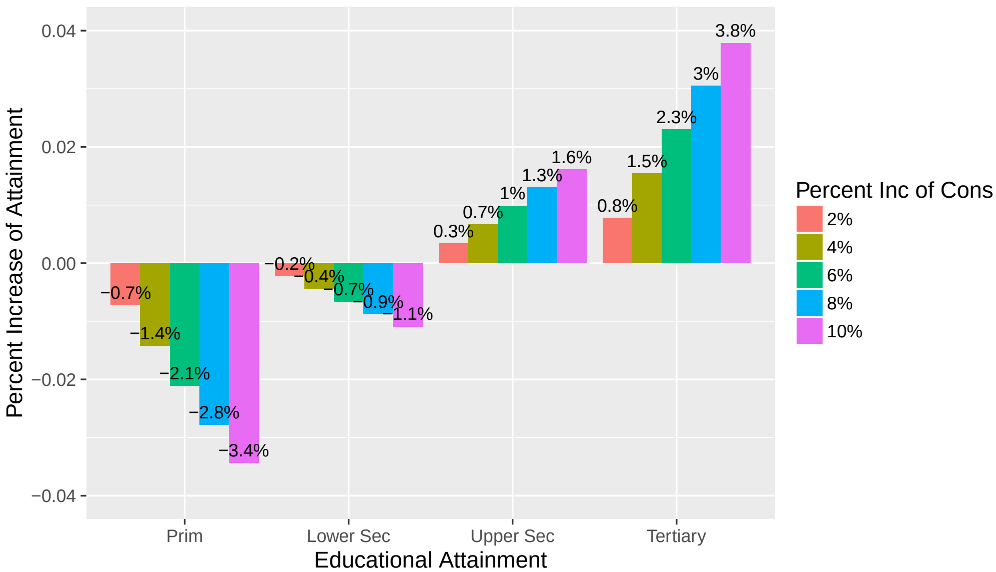
Since Logit has closed form probabilities, simulations are done using predicted probabilities of individuals and then by averaging these probabilities. This also corresponds to calculating population percentages for the final choices.

In the next four sections, two metrics, absolute and relative increase, will be used to define the change in educational attainment. Absolute increase in predicted probability could be explained with an example, e.g. policy causes tertiary education attainment from 10% to 11%, then absolute increase is 1%. Percentage increase in predicted probability could also be explained with an example, e.g. policy causes tertiary education attainment from 10% to 11%, then percentage increase is $10\% = (11 - 10)/10$.

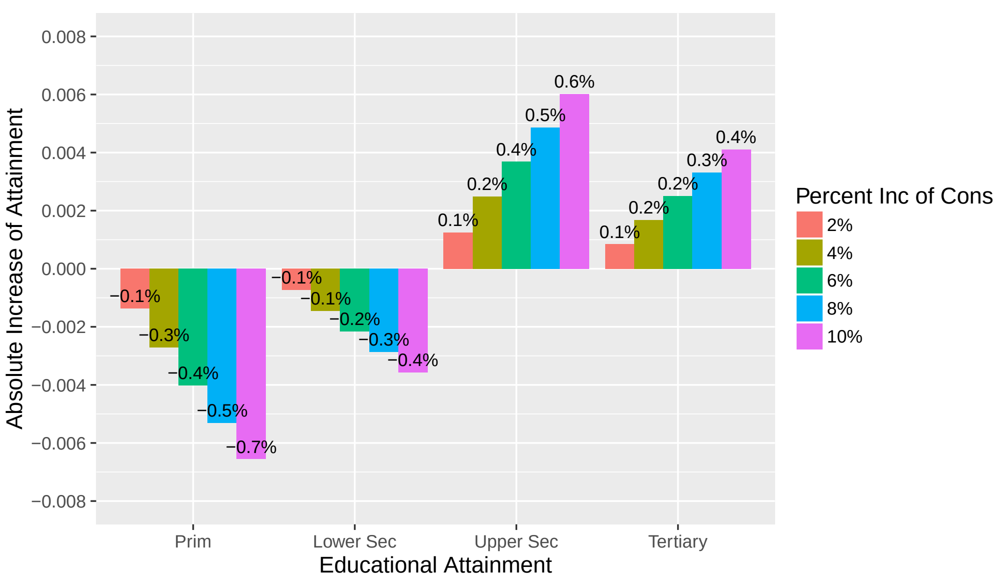
6.1.1 Consumption Increase

In this experiment, the goal is to understand the effect of economic growth in Indonesia, which is happening in the last few decades. In Section 6.1.3 I also analyze the effects of consumption increase for the low consumption demographic, which can be thought of as simulating effects of welfare programs.

Figure 6.1 simulates if per capita consumption is increased 2%, 4%, 6%, 8%, and 10% what would be the increase in the educational attainments. As can be seen that upper secondary and tertiary education are positively affected whereas primary and lower secondary show negative trend. The percentage increases are calculated within the conditional sets, i.e. percent increase in tertiary attainment is calculated with respect to the number of people who attained tertiary education. Also in Figure 6.1, the absolute increase is presented. Note that simulating an increase in consumption may not translate to a policy, i.e., fee subsidies. We may just be measuring what happens if the family is richer to start with or what happens with the effects of general economic growth. I include simulations with consumption increase throughout this section to give an intuition without claiming strong policy implications.



(a) Relative Increase



(b) Absolute Increase

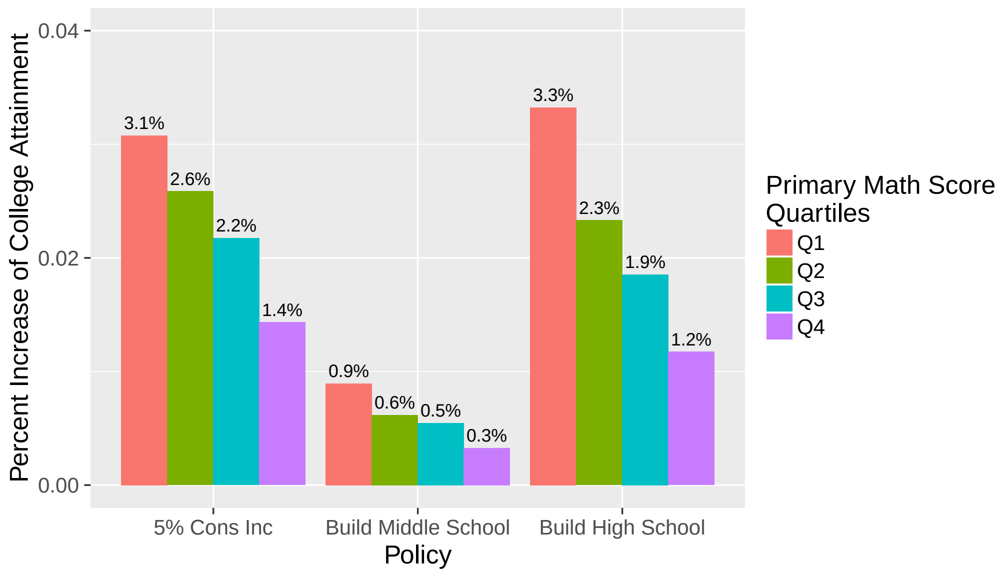
Figure 6.1: Simulating the effect of increasing per capita consumption

6.1.2 School Construction with respect to Math Score Quartiles

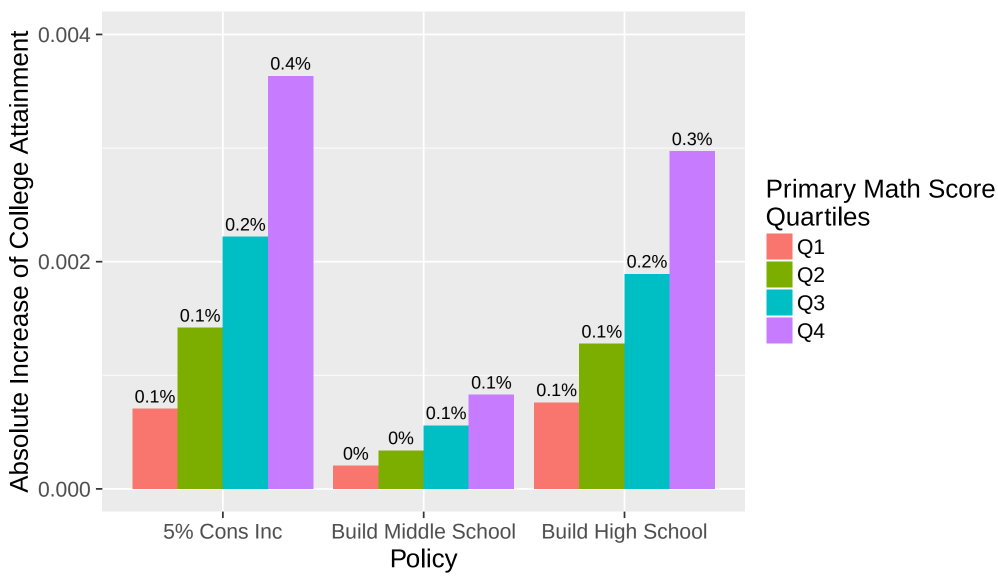
Figure 6.2 simulates three policies, -increase per capita consumption, -build a middle school, and -build an high school that students can access. Population is divided into four groups through quartiles of primary math scores. Table 6.1 shows descriptive statistics of these groups. It can be seen that there is no radical difference in the background characteristics of these groups but we see a big difference in their tertiary educational attainments. This is aligned with the fact that primary math score coefficients are quite significant in the third stage of the model.

Table 6.1: Descriptive statistics with respect to primary school math score quartiles.

Math Score Quartiles	Median Log Consumption Per Capita	Average Number of Middle Schools used by Community	Average number of High Schools used by Community	Tertiary Education Percentage
Q1	12.86	3.22	2.78	0.03
Q2	13.03	3.25	2.76	0.06
Q3	13.10	3.13	2.92	0.13
Q4	13.34	3.06	2.93	0.24



(a) Relative Increase



(b) Absolute Increase

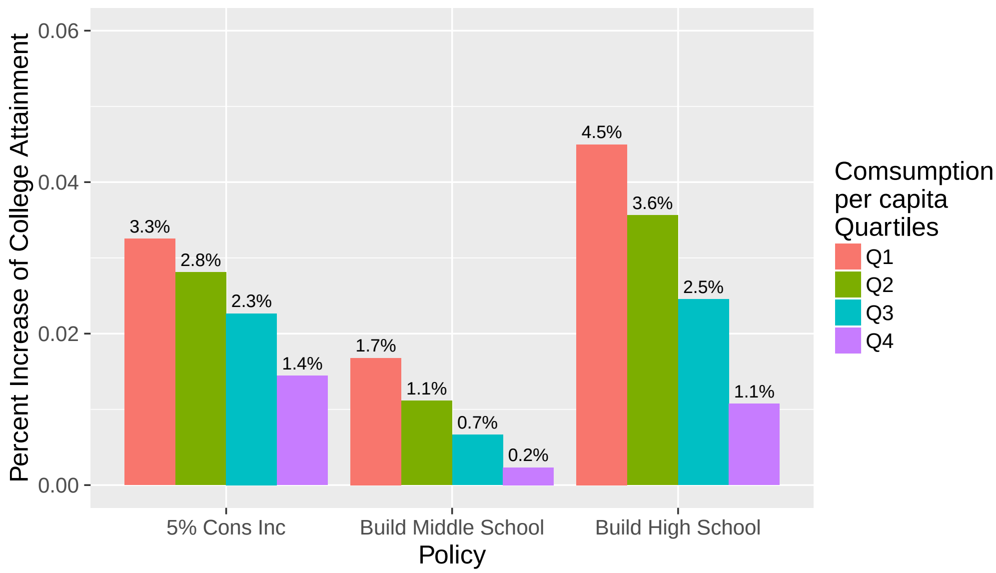
Figure 6.2: Comparison of policies of increasing consumption and building middle and high schools

6.1.3 School Construction with respect to Consumption Quartiles

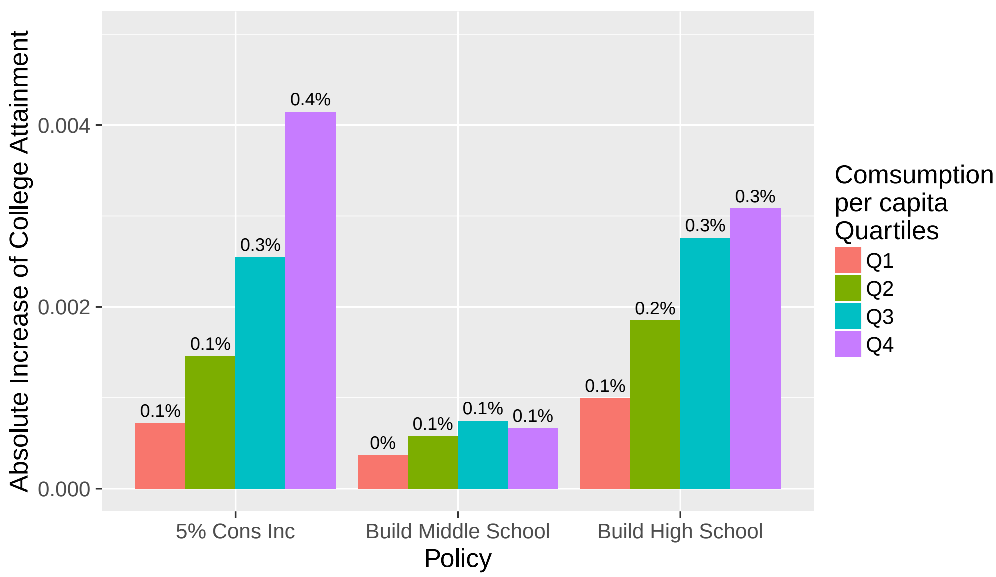
Figure 6.3 runs again a very similar simulation with a difference that the population groups are generated with respect to their consumption levels. More specifically, I compare people at different consumption quartiles to see how big the effect is for the very poor. As can be seen in Table 6.2, there is not a big difference between average number of middle schools and high schools for different consumption quartiles, however it could be seen that higher consumption level means more tertiary education. As can be seen in Figure 6.3, building an high school has a relative increase of 4.5% for lowest percentile, and the absolute change for the population is approximately 0.2%.

Table 6.2: Descriptive statistics of individuals with respect to consumption quartiles

Log Consumption	Average Number of Middle Schools	Average Number of High School	Tertiary Attendance
12.42	3.00	2.60	0.02
12.95	3.09	2.94	0.08
13.38	3.28	2.74	0.11
14.02	3.01	2.72	0.28



(a) Relative Increase



(b) Absolute Increase

Figure 6.3: Comparison of policies of increasing consumption and building middle and high schools

6.1.4 School Construction with respect to Urban or Rural Location

As can be seen in Table 6.3, urban areas has more schools than rural areas as expected. Here, I try to see what happens to tertiary attainment if I increase the number of high schools and middle schools by one in both locations. Figure 6.4 simulates three policies. It can be seen that there is 3.3% relative increase of tertiary attainment by building one more high school and 0.2% absolute increase.

Table 6.3: Descriptive statistics of individuals with respect to location

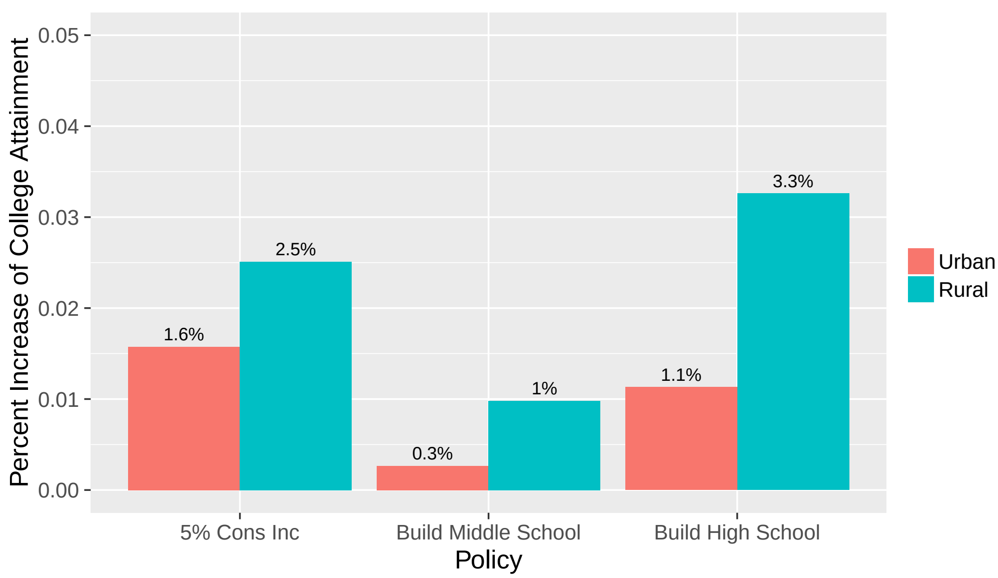
	Log Consumption	Average Number of Middle Schools	Average Number of High School	Tertiary Attendance
Urban	13.37	3.50	3.26	0.22
Rural	12.97	2.90	2.50	0.06

6.2 Policy Simulations on School Quality

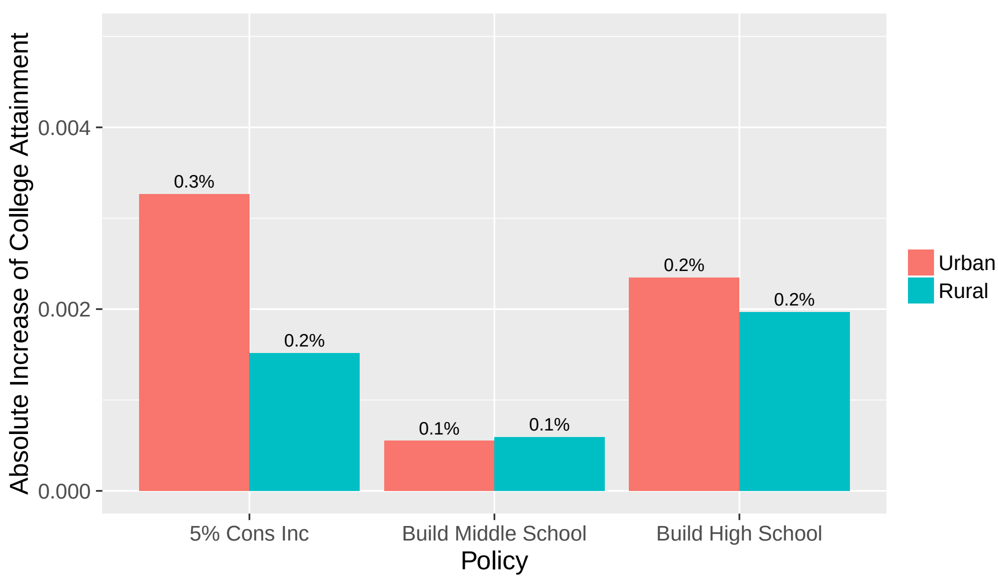
6.2.1 Java vs Non-Java Simulations

In this section, I show that improving primary school quality has a very big impact on tertiary attainment.

Java island is more populated and developed compared to the rest of the Indonesia. I also observe that school quality in Java Island is far better than



(a) Relative Increase



(b) Absolute Increase

Figure 6.4: Comparison of policies of increasing consumption and building middle and high schools

the rest. For policy purposes, I am asking a simple question: Would equating average community school quality of Non-Java islands to Java island yield an increase in tertiary education and how much?

As can be seen in Figure 6.5, I first equate the conditions in Non-Java Islands to Java island, in terms of labor market and job opportunities. This is done by adding Java fixed effects to non-Java students. Since living in Java island has a negative fixed effect on educational attainment, I consider this effect, ie. opportunity cost of living in Java island. Then I equate the number of middle schools and number of high schools in Non-Java islands to Java island, and then school quality and family background. This is done by calculating the difference in means and shifting everyone's covariates in non-Java by this difference. The key result is that primary school quality shift shows almost as big an attainment jump as the family background for tertiary attainment, which points the need for boosting the school quality. On the other hand, equating number of schools has minimal effect on tertiary attainment. This demonstrates that at this point in time, improving school quality would have a bigger impact.

6.2.2 School Quality of Indonesia in Each Province

School quality covariate used in my model is zero mean and has negative values by design. To make it easier to visualize, in the following graphs, I increment it by its standard deviation and call it "school quality index".

School quality varies from province to province as well as urban to rural.

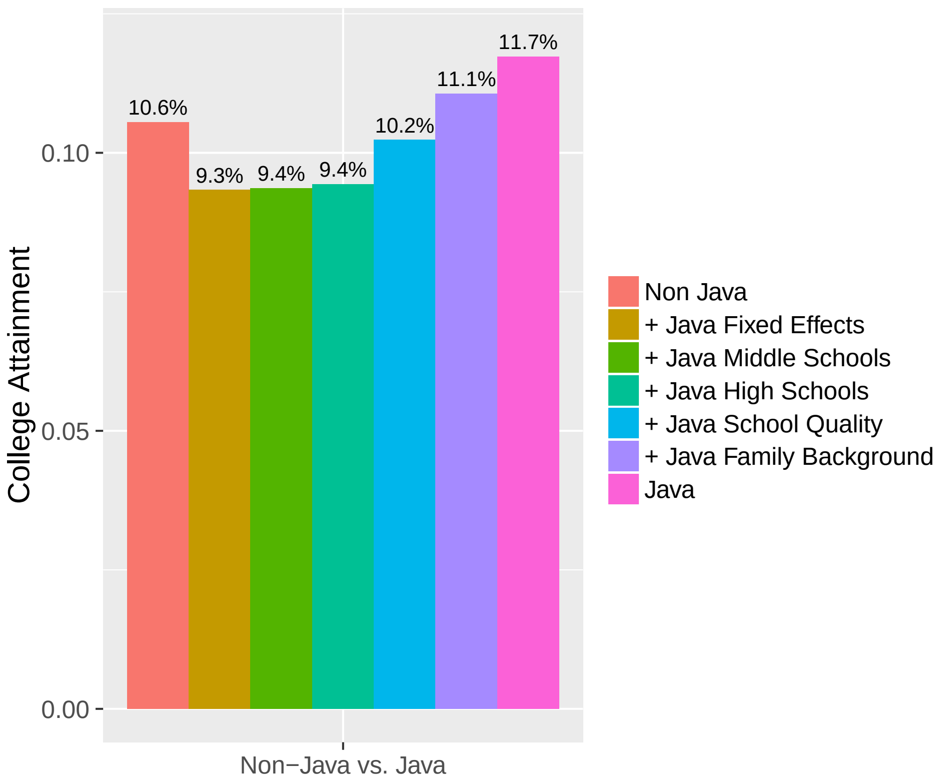


Figure 6.5: Non-Java vs Java Comparision

In Figure 6.6, the school quality of each province as well as of urban and rural sub-populations are displayed. In Figure 6.7 school quality by each province is presented on Indonesian map for better visualization. As can be seen from both figures, most quality schools are located in most populous provinces of Indonesia, ie. Java Island. Observation of the disparities in school quality among the different provinces shows that West-Java provides the best quality education. Another interesting observation is that North Sumatra's school quality in rural areas are higher than the school quality in urban areas. One argument for this observation is from a study by the Kantor Statistik (1982). According to them, the Gini ratio of North Sumatra in 1982 was 0.267. The study confirmed a comparatively small "relative inequality" as a whole province and differentiating between urban and rural areas yielded similar results as well (Barlow and Thee, 1988).

6.2.3 Extension of School Quality Policy

For future research, it may be possible to connect the school quality covariate with teacher characteristics. For now, I give some of the descriptive statistics from IFLS. Based on the previous review done by Glewwe et al. (2011), teachers characteristics, specifically how good teachers teach, are most important on student's academic achievement.

I focus on the following primary school teacher characteristics: teacher's age to proxy teachers experience, teacher has gone to university or not, did they use old (1976) curriculum or the new (1984) curriculum in class. To get

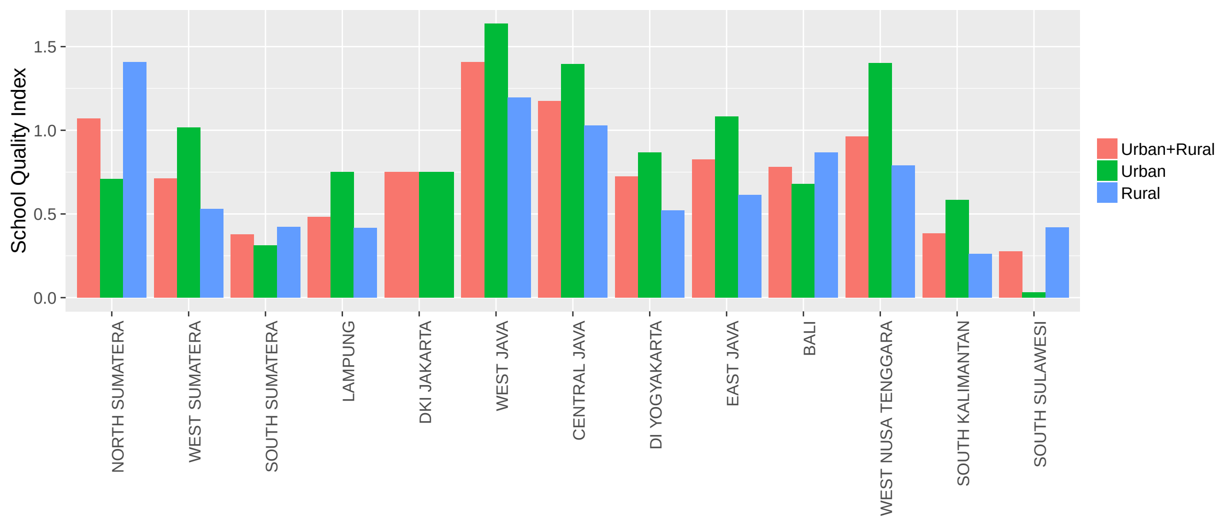


Figure 6.6: Province Primary School Quality

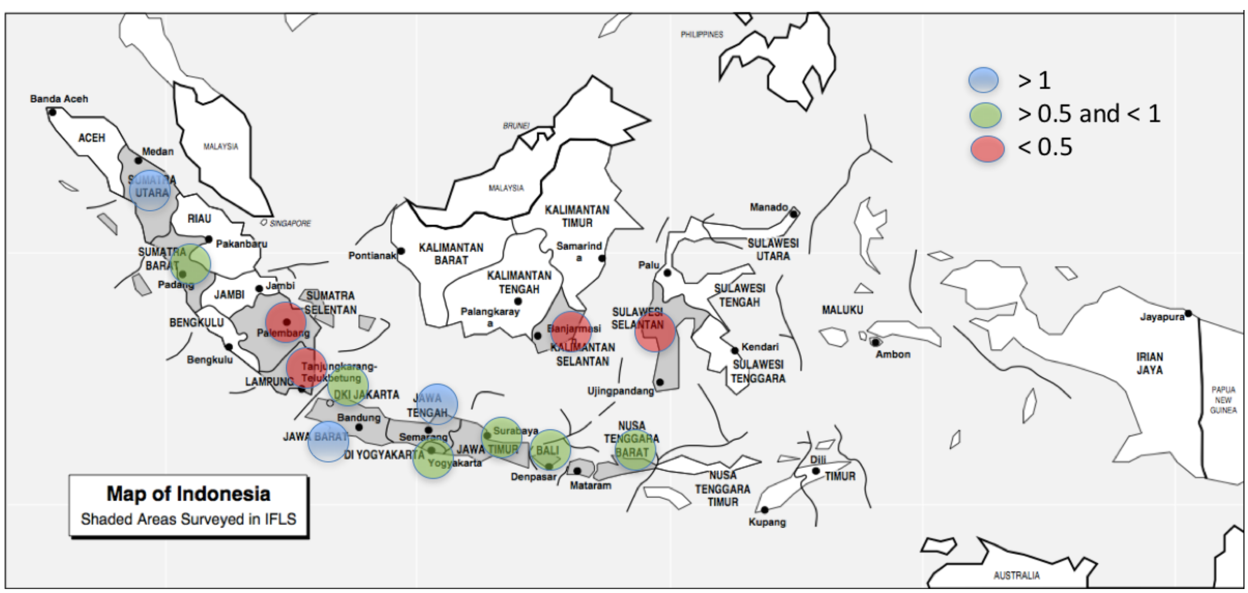


Figure 6.7: Province Primary School Quality Map

an idea, I regress community's primary school quality with these covariates to visualize the importance of these factors. As can be seen from the Table 6.4, teacher's age (how experienced a teacher is) affects the community school quality positively, whereas the usage of the old curriculum affects the community school quality negatively. The analysis below is just to shed a light on the relationship between the school quality and teacher's characteristics. Note that, more analyses as well as more covariates may be needed to investigate these issues in the future.

Table 6.4: Regression on community's school quality with community's average teacher age, teacher education and curriculum

	Estimate	Std. Error	t value	Pr(>|t|)	
Intercept	-0.5303	0.3197	-1.66	0.0984	.
Teacher age	0.0181	0.0085	2.14	0.0330	*
College or above educated	-0.2627	0.2387	-1.10	0.2722	
Old Curriculum	-0.4067	0.1936	-2.10	0.0367	*

Signif. codes: 0 *** 0.001 ** 0.01 * 0.05 . 0.1 1

Chapter 7

Conclusion and Future Work

7.1 Conclusion

In this dissertation, I analyzed the key factors of educational attainment, with a specific focus on tertiary attainment in a developing country. First, I analyzed the key factors affecting schooling decisions at each stage of the educational path, i.e., primary, lower secondary, upper secondary, and tertiary, since it is important to understand how the education inequities take place from the start. Then, I tried to unearth some casual relationships that could lead policies to encourage more high ability individuals to fulfill their potential and get university degrees.

To understand these education dynamics of Indonesia, I use the Indonesian Family Life Survey, which contains demographic, socioeconomic, and a detailed education information including EBTANAS (Indonesian National Exam) scores after each stage of students' educational path and Raven-Like test scores in order to measure their cognitive skills. IFLS Household survey data is also accompanied by detailed data about the communities and the facilities in these communities.

Since I try to analyze the key factors at each stage of educational path, a sequential choice model with unobserved heterogeneity is used, which handles the selection bias problem in education research.

7.1.1 Key Findings

Analyzing the model coefficients, it can be seen that parental background, region of residence, number of schools, and child's cognitive skills and non-cognitive personality traits are key factors of educational path decisions. Household per capita consumption, father's education, and mother's education are important in all stages of educational choices. Living in a rural area as opposed to an urban area negatively affects the earlier stages of educational path but does not seem to have an effect in the last stage. Being female has a negative effect on lower secondary school decision, and has a positive effect on tertiary decision. Birth order has a positive effect for second stage. Innate ability affects all stages of the educational path. Non-cognitive personality trait, i.e. smoking before age 15, has a negative effect for all stages.

After considering unobserved heterogeneity, coefficients of some of the covariates, such as living in the city at age 12, smoking before age 15, and log consumption per capita, are highly accentuated. At the third stage, most of the low ability students dropped out, so the coefficients stayed very close.

A striking result here is that fundamental math abilities acquired in early education may have a big impact on whether the student goes to university. This potentially demonstrates that there is a causal relationship between early education and bringing up highly educated individuals as well as most likely the lifetime utility of their earnings. Using math scores for a random sample of 25 students for each primary school surveyed, this research also attempts to understand the effects of primary school quality on educational attainment in Indonesia. I believe the key findings of this paper could shed some light on Indonesian's urgent problem of the lack of educated people at tertiary levels.

7.2 Future Work

> "The time will come when diligent research over long periods will bring to light things which now lie hidden. A single lifetime, even though entirely devoted to the sky, would not be enough for the investigation of so vast a subject..."
>
> *-Seneca, Natural Questions*

As a continuation of my research, more analysis could be done on specific school and teacher characteristics (teacher to pupil ratio, teacher's education, school infrastructure etc.), which appear to have strong positive impacts on upper secondary and university attendance.

Another IFLS wave is expected to be available in late 2016. With the availability of the new data set, future earnings can also be observed for my cohort. Thus, this new information may allow the use of wages as a future value for university decision for males. Also with the new wave, a new cohort could be added, who is under the effect of nine year compulsory education and also had a chance to finish their education fully by 2016. Since nine year compulsory education has been implemented in 1994 with the expectation to complete by the end of 2003/2004. Then, the effects of two different compulsory schooling regimes (6 year compulsory schooling and 9 year compulsory schooling with soft transition) in Indonesia can be compared. Especially with the new wave, there is a lot of possibilities to extend my work on education in Indonesia.

Bibliography

A Bruhin, T Epper, H FehrDuda (2010) "Risk and rationality: Uncovering heterogeneity in probability distortion," *Econometrica*, Vol. 78, pp. 1375–1412.

Almlund, M., A. Duckworth, J. Heckman, and T. Kautz (2011) "Personality psychology and economics.," *Handbook of the Economics of Education*, Vol. 4, p. 1181.

Amin, Samia, Jishnu Das, and Markus Goldstein (edited) (2007) "Are you Being Served: New Tools for Measuring Service Delivery," *World Bank, Washington DC*.

Arcidiacono, P (2005) "Affirmative Action in Higher Education: how do admission and financial aid rules affect future earnings?" *Econometrica*, Vol. 73, pp. 1477–1524.

Banerjee, Abhijit V., Shawn Cole, Esther Duflo, and Leigh Linden (2007) "Remedying Education: Evidence from two randomized experiments in India," *Quarterly Journal of Economics*, Vol. 122, No. 1235–1264.

Barber, Sarah, Sri Moertiningsih Adioetomo, Abdillah Ahsan, and Diahhadi Setyonaluri (2008) "Tobacco Economics in Indonesia," *International Union Against Tuberculosis and Lung Disease.*

Barlow, Colin and Kian Wie Thee (1988) "The North Sumatran Regional Economy: Growth with Unbalanced Development," *ASEAN Economic Research Unit, Institute of Southeast Asian Studies.*

Becker, Gary S. and H. Gregg Lewis (1973) "On the Interaction between the Quantity and Quality of Children," *Journal of Political Economy*, Vol. 81:2, No. 279–288.

Biglan, A. (2004) "Helping Adolescents at Risk: Prevention of Multiple Problem Behaviors.," *New York: Guilford Press.*

Bowles, S. and H. Gintis (2002) "The Inheritance of Inequality," *Journal of Economic Perspectives*, Vol. 16, pp. 3–30.

Buchmann, C (2000) "Family structure, parental perceptions and child labor in Kenya: What factors determine who is enrolled in school?" *Soc. Forces*, Vol. 78, No. 1349–1379.

Buchmann, C. and E. Hannum (2001) "Education and Stratification in Developing Countries: A Review of Theories and Research," *Annual Review of Sociology*, Vol. 27, No. 77–102.

Cameron, Lisa (2002) "The impact of the Indonesian financial crisis on children: Data from 100 villages survey," *Policy Research Working Paper. Washington DC: World Bank*, Vol. 1.

Cameron, S. V. and J. J. Heckman (1998) "Life cycle schooling and dynamic selection bias: Models and evidence for ve cohorts of American males.," *Journal of Political Economy*, Vol. 106, pp. 262–333.

——— (2001) "The dynamics of educational attainment for black, hispanic, and white males.," *Journal of Political Economy*, Vol. 109, pp. 455–499.

Cawley, J., J. J. Heckman, and E. J. Vytlacil (2001) "Three Observations on Wages and Measured Cognitive Ability.," *Labor Economics*, Vol. 8 (4), p. 419442.

Checchi, D. (2006) "The Economics of Education. Human Capital, Family Background and Inequality," *Cambridge University Press*.

Coleman, JS, Campbell E, Hobson C, Mc-Partland J, and Mood A (1966) "Equality of Educational Opportunity.," *Washington, DC: Dep. Health, Educ., Welfare*.

Cunha, F. and J. J. Heckman (2008) "Formulating, identifying and estimating the technology of cognitive and noncognitive skill formation," *Journal of Human Resources*, Vol. 43(4), p. 738782.

Dempster, A., N. Laird, and D. Rubin (1977) "Maximum Likelihood from Incomplete Data via the EM Algorithm," *Journal of the Royal Statistical Society*, Vol. 78, pp. 1–38.

Djutaharta, T and HV Surya (2003) "Economics of Tobacco Control Paper. Research on tobacco in Indonesia: an annotated bibliography and review on tobacco use, health effects, economics, and control effects.," *World Bank*, Vol. No:10.

Duflo, Esther (2001) "Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence from an Unusual Policy Experiment," *American Economic Review*, Vol. 91, No. 4, pp. 795–813.

Duncan, G. J., C. J. Dowsett , A. Claessens, K. Magnuson, A.C. Huston, P. Klebanov, L. Pagani, and C. Japel (2007) "School readiness and later achievement.," *Developmental Psychology*, Vol. 43, pp. 1428–1446.

Eckstein, Zvi and Itzhak Zilcha (1994) "The Effects of Compulsory Schooling on Growth Income Distribution and Welfare.," *Journal of Public Economics*, Vol. 54, No. 339–359.

Frankenberg, E. and L. Karoly (1995) "The 1993 Indonesian Family Life Survey: Overview and field report," *Publication no. DRU-1195/1-NICHD/AID. Santa Monica, CA: Rand Corporation.*

Frankenberg, E. and D. Thomas (2000) "The Indonesia Family Life Survey (IFLS): study design and results from waves 1 and 2," *Publication no. DRU-2238/NIA/NICHD. Santa Monica, CA: Rand Corporation*, Vol. 1.

Fuller, B (1987) "What school factors raise achievement in the developing world?" *Rev. Educ. Res.*, Vol. 57, No. 255–292.

Fuller, B and P Clarke (1994) "Raising school effects while ignoring culture?" *Rev. Educ. Res.*, Vol. 64, No. 119–157.

Gintis, Herbert (1971) "Education, Technology, and the Characteristics of Worker Productivity," *American Economic Review*, Vol. 61(2), p. 266279.

Glewwe, Paul, Michael Kremer, and Sylvie Moulin. (2009) "Many children left behind? Textbooks and test scores in Kenya," *American Economic Journal: Applied Economics*, Vol. 1, No. 112–135.

Glewwe, Paul, Eric A. Hanushek, Sarah D. Humpage, and Renato Ravina (2011) "School resources and educational outcomes in developing countries: A review of the literature from 1990 to 2010," *NBER Working Paper No. 17554*.

Gullone, E. and S. Moore (2000) "Adolescent risky-taking and the five-factor model of personality," *Journal of Adolescence*, Vol. 23, p. 393407.

Hanushek, Eric A., Victor Lavy, and Kohtaro Hitomi (2008) "Do Students Care about School Quality? Determinants of Dropout Behavior in Developing Countries.," *Journal of Human Capital*, Vol. 2(1), pp. 69–105.

Heady, Christopher (2003) "The Effect of Child Labor on Learning Achievement," *World Development*, Vol. 31(2), pp. 385–398.

Heckman, J., J. Stixrud, and S. Urzua (2006) "The Effects of Cognitive and Noncognitive abilities on Labor Market Outcomes and Social Behaviour," *NBER WP*, Vol. No. 12006.

Heckman, J., J. Humphries, G Veramendi, and S. Urzua (2011) "The Effects of Educational Choices on Labor Market, Health and Social Outcomes," *University of Chicago WP*, Vol. No 2011.

Heckman, James J., John Eric Humphries, and Gregory Veramendic (2016) "Dynamic treatment effects," *Journal of Econometrics*, Vol. 191(2), pp. 276–292.

Heckman, J. and S. Mosso (2014) "The Economics of Human Development and Social Mobility," *ZA DP*, Vol. No 8000.

Heckman, J. and B. Singer (1984) "A Method for Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration Data," *Econometrica*, Vol. 52, pp. 271–320.

Heyneman, SP (1976) "Influences on academic achievement: a comparison of results from Uganda and more industrialized societies," *Sociol Educ*, Vol. 49, No. 200–211.

Heyneman, SP and WA Loxley (1983) "The effect of primary school quality on academic achieve- ment across twenty-nine high- and low- income countries," *Am. J. Sociol.*, Vol. 88, pp. 1162–1194.

Jensen, A.R. (1987) “The g beyond factor analysis,” *The influence of cognitive psychology on testing*, pp. 87–142.

Jensen, Arthur (1998) “The g Factor: The Science of Mental Ability (Human Evolution, Behavior, and Intelligence).”

King, EM and LA Lillard (1987) “Education policy and school attainment in Malaysia and the Philippines,” *Econ. Educ. Rev.*, Vol. 6, No. 167–181.

Knight, John B. and R. H. Sabot (1990) “Education, Productivity, and Inequality: The East African Natural Experiment,” *Oxford:Oxford University Press*.

Lam, D., C. Ardington, N. Branson, K. Goostrey, and M. Leibbrandt (2013) “Credit constraints and the racial gap in post-secondary education in South Africa.,” *National Bureau of Economic Research*.

Lillard, Lee A. and Robert J. Willis (1994) “Intergenerational Educational Mobility: Effects of Family and State in Malaysia,” *The Journal of Human Resources*, Vol. 29, pp. 1126–1166.

Lockheed, ME and NT Longford (1991) “School effects on mathematics achievement gain in Thailand. In Schools, Classrooms and Pupils: International Studies of Schooling From a Multi-level Perspective,” *ed. SW Rau- denbush, JD Willms. San Diego: Academic*, No. 131–148.

Maralani, V. (2008) “The Changing Relationship Between Family Size and Educational Attainment Over the Course of Socioeconomic Development: Evidence From Indonesia.,” *Demography*, Vol. 45, No. 3, pp. 693–717.

Mare, Robert D. (1980) “Social Background and School Continuation Decisions,” *J. American Statistical Assoc.*, Vol. 75, pp. 295–305.

Metzler, Johannes and Ludger Woessmann (2010) “The impact of teacher subject knowledge on student achievement: Evidence from within-teacher within-student variation,” *IZA Discussion Paper*.

Montgomery, MR and CB Lloyd (1997) “Excess fertility, unintended births and childrens schooling,” *Policy Res. Div. Work. Pap., New York: Popul. Council*, Vol. 100.

Mullis, Ina V.S., Michael O. Martin, Pierre Foy, and Alka Arora (2011) “TIMSS 2011 International Results in Mathematics.”

Nawi, Ng, L. Weinehall, and A. Ohman (2006) “‘If I dont smoke, I’m not a real man’Indonesian teenage boys’ views about smoking,” *Oxford Journals Medicine and Health, Health Education Research*, Vol. 22(6), pp. 794–804.

Nichter, M., S. Padmawati, N Ng M. Danardono, and Y. Prabandari (2009) “Reading culture from tobacco advertisements in Indonesia,” *Tobacco Control*, Vol. 19, pp. 98–107.

OECD (2012) “PISA 2012 Results in Focus What 15-year-olds know and what they can do with what they know.”

Pal, Sarmistha (2004) "Child Schooling in Peru: Evidence from a Sequential Analysis of School Progression," *Journal of Population Economics*, Vol. 17, No. 4, pp. 657–680.

Parish, WL and RJ Willis (1993) "Daughters, education, and family budgets: Taiwan experiences," *J. Hum. Resour.*, Vol. 28, No. 863–898.

Radford, Neal and GE Hinton (1998) "A view of the EM algorithm that justifies incremental, sparse, and other variants.," *Learning in Graphical Models, Kluwer Academic Publishers*, pp. 355–358.

Riddell, AR (1989) "An alternative approach to he study of school effectiveness in third world countries," *Comp. Educ. Rev.*, Vol. 33, No. 481–497.

Riphahn, Regina and Guido Heineck (2009) "Intergenerational Transmission of Educational Attainment in Germany - The Last Five Decades," *Journal of Economics and Statistics*, Vol. 229, No. 36–60.

Rouse, C.E. and L. Barrow (2006) "U.S. elementary and secondary schools: equalizing opportunity or replicating the status quo?" *Future Child*, Vol. 16, pp. 99–123.

Siegler, Robert S., Greg J. Duncan, Pamela E. Davis-Kean, Kathryn Duckworth, Amy Claessens, Mimi Engel, Maria Ines Susperreguy, and Meichu Chen (2012) "Early Predictors of High School Mathematics Achievement," *Psychological Science*, Vol. 23(7), pp. 691–697.

Strauss, J., K. Beegle, B. Sikoki, A. Dwiyanto, Y. Herawati, and F. Witoelar (2004) “The third wave of the Indonesia Family Life Survey (IFLS3): Overview and field report,” *Publication no. WR-144/1-NIA/NICHD. Santa Monica, CA: Rand Corporation.*

Suryadarma, Daniel, Asep Suryahadi, Sudarno Sumarto, and F. Halsey Rogers (2006) “Improving student performance in public primary schools in developing countries: Evidence from Indonesia,” *Education Economics*, Vol. 14 (4), No. 401–429.

Thomas, Duncan, Kathleen Beegle, Elizabeth Frankenberg, John Strauss, Bondan Sikoki, , and Graciela Teruel (2004) “Education in a crisis,” *Journal of Development Economics*, Vol. 74, pp. 53–85.

Thomas, Duncan and Elizabeth Frankenberg (2001) “Lost But Not Forgotten: Attrition and Follow-up in the Indonesian Family Life Survey,” *Journal of Human Resources*, Vol. 36, No. 3, pp. 556–592.

Tomasi, Carlo“http://www.cse.psu.edu/ rtc12/CSE586/papers/emTomasiTutorial.pdf.”

Unesco (2010) “Secondary Education Regional Information Base: Country Profile Indonesia.”

Urquiola, Miguel (2006) “Identifying class size effects in developing countries: Evidence from rural Bolivia,” *Review of Economics and Statistics*, Vol. 88 (1), No. 171–177.

WDR (2007) “World Development Report.”

WHO (2011) "World Health Organisation Report on the Global Tobacco Epidemic."

WorldBank (1989) "Indonesia: Basic Education Study," *Report No 78414*.

——— (2014) "Indonesia Economic Quarterly : Hard Choices," *Report No 91560*.

Appendices

Appendix A

The derivation of EM Algorithm

Discussion in this section is continuation of Section 5.1.3 and closely follows tutorial from Tomasi.

As can be seen in Figure 5.2, at the E step, starting with initial parameters $\theta^{(t)}$, EM algorithm forms a lower bound with $g_{(t)}$ which is the objective function of $logP(x, \theta)$. At the M step, $\theta^{(t+1)}$ is computed as the maximum of $g_{(t)}$. At the next E step, a new lower bound $g_{(t+1)}$ is formed. At the next M step, $\theta^{(t+2)}$ is computed as the maximum of $g_{(t+1)}$.

Here is a short derivation of the EM algorithm based on the idea of the bound optimization[1]. For any probability distribution $Q(z)$:

[1] You could find the detail derivations in (Radford and Hinton, 1998)

$$log\left(\sum_z P(x,z;\theta\right) = log\left(\sum_z Q(z)\frac{P(x,z;\theta)}{Q(z)}\right) \geqslant \sum_z Q(z) log\left(\frac{P(x,z;\theta)}{Q(z)}\right) \tag{A.1}$$

The inequality holds with equality $Q(z) = P(z|x;\theta)$. The derivation in equation A.1 is based on Jensen's inequality for concave function.

$\hat{\theta}^{(t+1)} = \arg\max_\theta g_t(\theta)$ where

$$g_t(\theta) = \sum_z p\left(z|x;\theta^{(t)}\right) log\left(\frac{P(x,z;\theta)}{P(z|x;\hat{\theta}^{(t)})}\right) \tag{A.2}$$

Applying Jensen's Inequality, $g_t(\hat{\theta}^{(t)}) = logP(x;\hat{\theta}^{(t)})$.

Observe that, $g_t(\hat{\theta})^{(t)} \leqslant g_t(\hat{\theta})^{(t+1)}$ by definition of the update rule. Furthermore, $g_t(\hat{\theta})^{(t+1)} \leqslant logP(x;\hat{\theta}^{(t+1)})$ guarantees that $g_t(\theta)$ is a lower bound on $logP(x;\theta)$ for any parameter θ. The update rule results in monotonic improvement of the maximum likelihood objective for incomplete data.

To see the connection between equation A.2 and the description of the EM algorithm given in the text, consider the following equivalent update rule:

$$\hat{\theta}^{(t+1)} = \arg\max_\theta \sum_z P(z|x;\hat{\theta}^{(t)}) logP(x,z;Q) \tag{A.3}$$

Note the additive separability of this objective function. The fact that the objective function of A.3 differs from by a constant offset which does not depend on θ. In this final form, we see that the EM update rule effectively maximizes the log-likelihood of a dataset expanded to contain all possible completions z of the unobserved variables, where each completion is weighted by the posterior probability, $P(z|x:\hat{\theta}^{(t)})$.