Charles University in Prague

Faculty of Social Sciences
Institute of Economic Studies

MASTER THESIS

Effects of human capital in developing countries

Author: **Bc. Jan Mouček**

Supervisor: Julie Chytilová Ph.D.

Academic Year: **2014/2015**

Declaration of Authorship

The author hereby declares that he compiled this thesis independently and using only the listed resources and literature, and the thesis has not been used to obtain a different or the same degree.

The author grants to Charles University permission to reproduce and to distribute copies of this thesis document in whole or in part.

Prague, January 5, 2015

Signature

Acknowledgments

The author is grateful to Mrs. Julie Chytilová for supervision of this thesis and support in finding and researching the topic. The author is also grateful to Mr. Havránek and Mr. Mejstřík for oversight of the proceedings of this thesis and helpful tips.

Abstract

Improving education and health of people living in developing countries is an often discussed topic. We however know relatively little about the effects that investments into schooling and improving the nutrition and hygiene of people in those countries have on their productivity at work. This thesis investigates the effects of four variables representing investment into human capital on income of individuals from Bangladesh and Indonesia. The four studied variables are education, migration, height and Body Mass Index. The thesis aims at comparing results of estimation using instrumental variables estimator with previous similar research. It tries to observe possible differences in size of effects the human capital variables have. It finds however that the instrumental variables used are of insufficient quality for the estimation to yield reliable results. Because the instruments used were almost the same as in underlying works and their quality has not been previously tested the conclusion of this thesis is that further search for correct instrumental variables describing the human capital variables in various countries around the globe is necessary.

JEL Classification	C21, C36, E22, E23, I15, I25, J24, O15
Keywords	Human capital, education, health, height, BMI, migration, income determinants, endogenous variables, instrumental variables

Author's e-mail	jan.moucek@gmail.com
Supervisor's e-mail	chytilova@fsv.cuni.cz

Abstrakt

Zvýšení kvality vzdělání a zdraví lidí žijících v rozvojových zemích je často diskutovaným tématem. Přesto víme relativně málo o efektech investic do školství a zkvalitnění výživy a hygieny lidí z těchto zemí na jejich produktivitu práce. Tato práce zkoumá efekty čtyř proměnných, zastupujících investice do lidského kapitálu, na příjem osob z Bangladéše a Indonésie. Tyto čtyři studované proměnné jsou vzdělání, migrace, výška a Body Mass Index. Tato práce se zaměřuje na porovnání výsledků odhadu s využitím instrumentálních proměnných s předchozí prací v tomto oboru. Snaží se vysledovat možné rozdíly mezi výší efektů, které proměnné lidského kapitálu mají v různých zemích. Ukazuje však, že použité instrumentální proměnné nemají dostatečnou kvalitu, aby odhady přinesly spolehlivé závěry. Protože použité instrumenty byly téměř totožné s podkladovou prací a jejich kvalita nebyla dříve testována, závěr této práce je, že je nezbytné dál hledat vhodné instrumentální proměnné vysvětlující lidský kapitál v různých zemích světa.

Klasifikace	C21, C36, E22, E23, I15, I25, J24, O15
Klíčová slova	Lidský kapitál, vzdělání, zdraví, výška, BMI, migrace, determinanty příjmu, endogenní proměnné, instrumentální proměnné

E-mail autora	jan.moucek@gmail.com
E-mail vedoucího práce	chytilova@fsv.cuni.cz

Contents

List of Tables

Master Thesis Proposal

Institute of Economic Studies
Faculty of Social Sciences
Charles University in Prague

Author:	**Bc. Jan Mouček**	Supervisor:	**Julie Chytilová Ph.D.**
E-mail:	jan.moucek@gmail.com	E-mail:	chytilova@fsv.cuni.cz
Phone:	728539698	Phone:	222112322
Specialization:	Finance, Financial Markets and Banking	Defense Planned:	January 2015

Proposed Topic:

Effects of human capital in developing countries

Motivation:

Human capital influences productivity of workers in a very important way. Unfortunately while investments into factory machines bring clearly measurable profits and even investments into R&D have measureable effects on innovation, investments in human capital such as education or health are hard to measure. That makes it difficult to decide about policies concerning these investments.

Investments are most needed in developing countries and investments in human capital in these countries are often called for from public. These investments are an interesting topic for people living in developed countries as well since a lot of these investments are done not by government but by charities sponsored by their money. Possible information regarding effects of these investments could therefore change the thinking of both policymakers of developing countries, but also thinking of charity project managers and donors.

My work is strongly affected by paper by Schultz (2003). Schultz looks for effects of human capital on wage in Ghana and Ivory Coast finding endogeneity as expected in at least some types of human capital. He solves this endogeneity by instrumental variables and then further studies cross-effects of human capital and mentions differences between these effects between men and women and in between Ghana and the Ivory Coast, his explanation of these differences is that it's given by different levels of poverty and infrastructure, nevertheless he doesn't go too deep into other possibilities or other effects. I would like to replicate works of Schultz on data from other countries using same or similar methods and same instrumental variables for human capital. I would like to focus on reasons behind differences in the effects between countries.

Hypotheses:

1. Hypothesis #2: Effects of human capital on wage are not equal in different countries
2. Hypothesis #3: Size of effects of human capital are associated with level of GDP
3. Hypothesis #4: Size of effects of human capital are associated with level of poverty

Methodology:

I will be using data obtained from rand.org from household surveys in Indonesia and Bagladesh from 2007 and 1995 respectively. Together with the results of Schultz (2003) from Ghana and Ivory coast these should give me more insight into reasons behind different levels of effects of human capital on worker productivity. I will be using the aforementioned data because they enable me to use same or similar variables as Schultz is using in his work. This condition is difficult to hold for Schultz uses many very specific instrumental variables in his research, but is necessary to hold for otherwise the results of my research and comparison of my and Schultz's results would be flawed.

I am going to use econometric models based on OLS and instrumental variables similarly to Schultz (2003) to measure effects of human capital on income. I will use the same instrumental variables as in aforementioned paper by Schultz to avoid using wrong IVs. I will look for obvious correlation between size of effects of human capital on wage and levels of GDP, poverty, infrastructure or geographical effects and try to further confirm (or reject) Schultz's claims about poverty being the main reason for differences in effects of human capital in different countries. Schultz gives no other explanation to difference in effects of human capital in his two samples than general wealth of country and level of poverty. I will be focusing on confirming this claim that there is larger payoff for being strong and healthy in poorer countries. I will also study cross effects of human capital, again similarly to Schultz (2003).

Expected Contribution:

I plan to investigate differences in effects of human capital in developing countries across the globe. In contrast to previous literature in the field I plan to address larger amount of data in more countries and more contrasting environments (according to World Bank (actual) GDP per capita in Indonesia is about twice as big as in Ghana and about four times bigger than in Bangladesh, not mentioning its vast political, infrastructural and foremost cultural differences). This will enable me to give more precise explanation then previously assessed as to reasons behind the differences in size of effects of particular type of human capital. This could give policy makers and charity project planners hints as to where should help be addressed in developing countries in order to improve productivity of local workforce.

Outline:

1) Introduction
2) Literature overview
3) Data
   a) Source of data
   b) Description of data
   c) Possible issues
4) Model specification
5) Estimates of human capital effects
6) Reasoning differences in human capital effects
7) Conclusion

Core Bibliography:

Becker, G.S., 1993. Human Capital. Third Edition, Chicago: University of Chicago Press

Behrman, J.R., 1993. The Economic Rationale for Investing in Nutrition in Developing Countries. World Development 21, 1749-1771

Fogel, R.W., 1994. Economic Growth, Population Theory and Physiology. American Economic Review 84, pp. 369-395

Schultz, T.P., 2003. Wage Rentals for Reproducible Human Capital: Evidence from Ghana and the Ivory Coast. Economics & Human Biology, Elsevier, vol. 1(3), pages 331-366, December

Strauss, J., Thomas, D., 1995. Human recourses: Empirical Modeling of Household and Family Decisions. In: Behrman, J.R., Srinivasan, T.N. (Eds.), Handbook of Development Economics, Vol. IIIA, North-Holland Pub. Co., Amsterdam, Chap. 34

Author **Supervisor**

1 Introduction

To help developing countries is the main goal of many institutions, charitable organizations and even life goal to some people. It is a very worthy cause. But how do we help the developing world efficiently? The opinions on what is the best way to tackle the gap in economical power between developing and developed world differ substantially across the population. Perhaps one of the views could be to improve labor productivity of developing countries through improving the health and education of their population. And truly many charitable organizations focus on improving health and educational standards in these countries. Obviously the main reason behind these efforts may not be to affect productivity of workers directly but more generally improving the quality of life of people in these countries. Still, from the perspective of an economist it seems interesting to look at the effects such efforts may have on productivity.

And looking at these effects is the objective of this thesis. It aims to explore the effects human capital has on wage and therefore, according to economic science, also on productivity. This thesis looks at the effects of height and weight which we assume to be given by health and quality of nutrition in both long and short term. This makes them representative of health investments into human capital. It looks at effects of education in terms of number of years spent in school. It also looks at the effects of migration as a form of investment into one's own human capital by moving residence to an area with better employment opportunities or with a potential for higher incomes. More importantly it tries to approach the reasons behind the differences in these effects among different countries. Why is it that in one country the productivity gain from one year of education is higher than in another country? This is an example of a question this thesis tries to answer.

Methodologically this thesis follows the work of Schultz (2003) to find out effects of human capital on wages in Indonesia and Bangladesh. We follow this work as precisely as possible to make comparison of the results relevant. This procedure includes foremost statistical estimates of effects of human capital in terms of height, BMI, education and migration on wages using optimal least square estimator and two-stage least square estimator. Furthermore we try to test quality of instruments explaining human capital used by Schultz. We try to base some limited conclusions about possible reasons behind differences in human capital effects on analysis of ours and Schultz's results.

The thesis is structured in the following way: Following this Introduction Chapter two gives brief overview of literature connected to the subject of human capital and its effects on productivity including overview of previous empirical research in the area. Chapter three describes the theoretical framework that we are using to approach the topic and introduces the statistical instruments we used to obtain the estimates. Chapter four gives description of the data used in the statistical models and basic macroeconomic data used to compare the studied countries. It also features general political and demographic overview of studied countries with emphasis on details that may affect behavior of local inhabitants in a way that may have impact on our results. Possible issues with our data are also discussed in this chapter. Chapter five than discusses the estimates obtained from statistical modeling. Both results of single equation and multiple equation models are included in Chapter five. Chapter six is dedicated to testing quality of instrumental variables used in this thesis and therefore also Schultz's (2003) paper. Chapter seven compares the results obtained from our models with the results of Schultz (2003) and discusses possible reasons behind differences in effects of human capital within the results across various countries. Conclusion and notes on possible future research are included in Chapter eight.

2 Literature Overview

In this chapter we want to shortly overview the literature that has been dedicated to investigating human capital in developing countries so far. We believe human capital effects in developed countries have been studied in much more detail and there is much interesting literature on the topic. For example Baum and Ford (2004) or Case and Paxson (2008) study the physical attributes and their effects on labor market discrimination. But the links between human capital and income are different in developed world compared to the developing countries, especially the effects of body structure as measured by BMI is often the very opposite at the measured levels. For these reasons literature from developed countries is only of limited value to our research.

The research of human capital is sometimes dived into two groups of human capital: educational human capital and health human capital. We nevertheless study both aspects. E.g Griliches (1977) discusses the effects of schooling on income and econometric issues relating to its estimation. Overview about the empirically assessed returns on education on global scale has been presented in Psacharopoulos (1994). Schultz (1997) shows that investment into better nutrition increases individuals productivity, which in return spurs economic growth and that way lowers poverty. Even before that Steckel (1995) documented the effects of stature and mostly height on income and individual welfare even with respect to historical data. Thomas and Frankenberg (2002) further study the effects that health in general has on economic prosperity.

These effects have been often empirically studied on data from farmers. Strauss (1986) looks for effects of intake of calories on productivity of farmers using data from Sierra Leone finding a clear link between nutrition and productivity. Deolalikar (1988) checks for both caloric intake and BMI effects on farmer's production in rural India. His finding do not find a clear link between daily productivity and intake of calories, but finds a link with BMI suggesting that good nutrition does not have immediate effect, although a short term effect is evident. Haddad and Bouis (1991) have studied effects of nutrition on farm-output in rural parts of southern Philippines with similar results. Thomas and Strauss (1997) report on effects of health on inhabitants of urban Brazil again finding a link between Body

Mass Index and wages. This is an important research for finding positive effects of health and body structure also in urban areas.

Other than purely econometric examination, field experiments were also conducted. For example Thomas et al. (2006) were studying effects of treatment with iron on workers in Central Java in Indonesia. They found a clear effect on the treated men raising their productivity and economic success while lowering their health problems. They found similar results only slightly diminished on women.

The combined approach of accepting both health and schooling as important variables for income in developing countries is documented by Strauss and Thomas (1995). Similar research to ours is presented by Kedir (2013). Kedir (2013) uses panel data from Ethiopia to find effect of height, BMI and education on wages. The author controls for endogeneity of education and BMI and finds positive results. These results show a significant effect of height for men's wage, but not for women's wage. They also show that education is more beneficial to women than to men. Schultz (1982) also mentions the importance of migration for income and the investment into migration as a factor to be taken seriously.

Our research is heavily based on the paper of Mr. T. P. Schultz from year 2003 called 'Wage rentals for reproducible human capital: evidence from Ghana and the Ivory Coast' as presented in the journal Economics & Human Biology. Since we are trying to replicate the research presented in this paper as closely as possible we deem it reasonable to make a short summary of the structure of the aforementioned paper. Schultz (2003) studies the data from Living Standards Measurement Surveys conducted between 1985 and 1989 in the Ivory Coast and Ghana. He uses four forms of human capital: height for indicating childhood nutrition, education in years of schooling, Body Mass Index to indicate adult nutritional status and current health, and migration from the region of birth. The author shows pairwise correlation between the individual human capital variables and then presents his Optimal Least Square estimates in form of gradually adding further variables to show omitted variable bias. Schultz (2003) admits endogeneity of human capital variables and tests it by Hausman-Wu endogeneity test finding only height and BMI to be endogenous and leaving education and migration exogenous. Sets of IV estimations with assuming both all of the variables and only those confirmed to be endogenous by testing are present in the paper. He finds stronger effect of education for men than for women and stronger effects in the Ivory Coast than in Ghana. Effects of the remaining forms of human capital are in sizes rather specific to the dataset. Schultz (2003) explains the difference in results mostly by the different macroeconomic wealth (GDP and poverty ratios) between the countries. Our thesis is focusing on this claim and trying to find further clues to support or deny it. Schultz (2003) further expands on his modeling by allowing for the human capital in its quadratic form and for substitute and complementary effects of human capital variables. He lastly confirms his results on a smaller sample of panel data from the Ivory Coast.

3 Human capital and income

This chapter describes the theoretical connection between human capital and income. Description of statistical framework employed to achieve estimates of effects of human capital on income and details on various statistical tools used is also included in this chapter.

3.1 Ordinary least square estimation

Throughout this thesis we use semi-logarithmic linear approximation of income function where income (W) is dependent on vector of human capital variables (I) and vector of other income-affecting exogenous variables (Y). The function looks as follows:

$$\log(w_i) = \beta_0 + \sum_{j=1}^{4} \gamma_j \, I_{ij} + \delta Y_i + \varepsilon_i \qquad (1)$$

For the first part of our modeling we will presume that the human capital variables are exogenous. We will test this exogeneity, which means that $cov(\varepsilon_i, I_{ij}) = 0$ for $j = 1, \dots, 4$ later by using standard Hausman-Wu test.

We estimate equation (1) with the universally recognized estimator – the Ordinary least square estimator (OLS). The Stata 12 statistical application was used for all of the calculations and estimations throughout this entire thesis.

3.1.1 Non-linear form

It may also happen that our model is a little too simplistic. There may be reasons to believe that there are diminishing returns to human capital. For example it is plausible that increase in incomes may be higher with same gains in health and physical growth for extremely malnourished population. Also Psacharopoulos (1994) reports that return to additional year of schooling are smaller at higher levels of education.

Furthermore our human capital variables may very well work as complements or substitutes to each other. For example an educated bureaucratic employee may not have much use of more physical strength. Emphasis here is on the fact that he is educated, not that he has a job where he can use it. It is therefore also the cross-effects of human capital that we should care for.

For these reasons we will also use a more flexible form of log-income function allowing for nonlinearity and considering cross-effects of human capital variables.

$$\log(w_i) = \beta_0 + \sum_{j=1}^{4} \gamma_j I_{ij} + \sum_{j=1}^{4} \sum_{k=j+1}^{4} \eta_{jk} I_{ij} I_{ik} + \sum_{j=1}^{3} \theta_j I_{ij}^2 + \delta y_i + \varepsilon_i \qquad (2)$$

Please note that the square effect of migration is not estimated, since we use a binominal variable representing it with value 1 for people that have migrated from their place of birth and 0 for people who stayed in the same area they were born in.

3.1.2 Endogeneity

There are several possible sources of endogeneity of human capital variables in our income function. One possible reason for human capital to be endogenous in the income function is that there are some unobserved exogenous differences across individuals that make them invest in human capital in either complementary or compensatory manner. One example of this could be “ability” of people. By ability we understand a genetic predisposition to achieve better results than others with equal investment or achieving same results with less investment. For example it may be that able people invest more into education because they get more return from it. Or maybe the other way around able people invest less into education because their ability enables them to gain same amount of knowledge with actually less education and they get to a point of diminishing marginal returns sooner than less able people (see Willis and Rosen, 1978). Since these effects are unobserved they are contained in our error term in the income function, so existence of such effects would mean correlation between the error term and the human capital variables.

3.2 Instrumental variables model

To eliminate possible endogeneity of human capital variables, which was discussed above, we also estimate the income function using instrumental variables estimator. In particular we use the Two-stage least square method (2SLS). This method constructs five equations: one explaining income and one for each of our human capital variables.

The first equation looks the same as the one we use in the OLS variant described in chapter 3.1. It describes income using our human capital variables and a vector of exogenous variables. The other four equations employ a new vector of so called instrumental variables (X). These variables are specifically chosen to be uncorrelated with the error term of the first equation, but to explain the human capital variables. As an example we can mention parent education. For example we assume that education of one's parents does not directly affect ones income as it is not an important thing for an employer to decide on wage increase or job acceptance. But it is believed that more educated parents lead their children to be also more educated therefore parent education directly affects child's education.

The equations look as follows:
The equation explaining logarithm of income:

$$\log\ (w_i) = \beta_0 + \sum_{j=1}^{4} \gamma_j\, I_{ij} + \delta Y_i + \varepsilon_i \tag{1}$$

And four equations explaining human capital variables

$$I_{ij} = \beta_0 + \delta Y_i + \alpha X_i + \varepsilon_i \tag{3}$$

Using 2SLS means that we firstly estimate the human capital variables using their respective equations via OLS (First stage). And then use the fitted values of human capital obtained this way to estimate the main (income) equation (Second stage).

We test the human capital variables for endogeneity separately as well as collectively. Schultz (2003) only finds endogeneity for some of his human capital variables. He afterwards runs a separate estimation where he only assumes endogeneity for those variables found to be endogenous by his testing. We proceed the same way.

3.2.1 Testing quality of instrumental variables

The instrumental variables we use in our IV estimations are as similar to the ones used by Schultz (2003) as possible to enable for a comparison of results. Nevertheless Schultz does not test the quality of his instruments as part of his paper. There are two possible issues with instruments of low quality. They generally belong to the two requirements we have on instrumental variables.

One of them is overidentification. This problem occurs when instruments do explain the endogenous variable, but they are at the same time correlated to the error term of the structural equation – in our case the log (income) equation. This points out that these variables would better be used in the structural equations and should not be used as instrumental variables for endogenous variables.

Second problem occurring are weak instrumental variables. Weak instruments do not explain the endogenous variables – in our case human capital – well enough. In other words their explanatory power is too low. The instrumental variable estimators do not work well with weak instruments and their estimations are unreliable. We test both of the aforementioned problems with results of the testing presented in Chapter 6.

4 Data

This Chapter gives basic economical and political information about the countries in question and produces overview of the microeconomic data used in the models. Throughout this chapter, possible issues connected to this data, are also discussed.

4.1 Country overview

This thesis uses data from two Asian developing countries: Indonesia and Bangladesh. This is the first difference from the original work by Schultz (2003) that we are trying to reproduce. Schultz (2003) uses data from two African developing countries: the Ivory Coast and Ghana. To understand the reasons behind differences in returns on investment into human capital it is necessary to look at some of the main demographic and political specifics of these countries.

4.1.1 Indonesia

Indonesia is an island country located in Southeast Asia. It is situated on five large islands: Sumatra, Java, Borneo, Sulawesi and Irian Jaya and over thirteen thousand smaller islands. It is the fourth largest country by population in Asia after China, India and Saudi Arabia. The population was estimated at over 221 million inhabitants in 2005 (UN as cited by Worldmark Encyclopedia of Nations, 2007). Almost 90% of population is Muslim. At the time our data was taken the government of Indonesia was attempting to lower Indonesian fertility rates to slow down its population growth. Resettlement of people from over-populated areas to less populous islands is another long-term governmental policy possibly affecting our data. Notable violent outbreaks and riots were happening in second half of 1990s connected to attempts at separation of East Timor from Indonesia and described as riots between Catholics and Muslims. Indonesia is also notably often stricken by earthquakes and tsunamis. A large tsunami prior to our data collection happened in 2004 and estimated 160,000 people were killed or missing because of it and it caused estimated $4.5 billion worth of damage. (Worldmark Encyclopedia of Nations, 2007)

Although Indonesia is rich in natural resources, especially oil and coal, most Indonesian people still work in agriculture with rice being the main crop harvested. Tourism is also important part of Indonesian economy as it accounted for 40% of GDP in 2005, even with a sharp drop after terrorist attacks on Bali in 2002. Until 1999 the country was run autocratically by President Suharto, using many tools of centrally planned economies. The country was experiencing an economy boom until the Asian crisis in 1997 that struck Indonesia very sharply. Nevertheless after few years of both economical and political downfall Indonesian economy was again regaining power in the early 2000s. In 2004 adult literacy was estimated at over 90% for male population and over 80% for female. Education is free and compulsory for first 9 years with first 6 years being covered by primary schools and next 3 years by junior secondary schools. Students may then choose to continue another 3 years of secondary schools in either general studies, Islamic studies or vocational studies. There is no national minimum wage. Minimum wage is set regionally by local governments however these rules are often not adhered to. (Worldmark Encyclopedia of Nations, 2007)

4.1.2 Bangladesh

Bangladesh is situated in South Asia on coast of Bay of Bengal neighboring India and Myanmar (Burma). It's less than one tenth of the size of Indonesia in terms of area. The population of Bangladesh was estimated in 2005 at 144 million inhabitants. Similarly to Indonesia, Bangladesh is fighting its overpopulation problem and local government has carried out a program to reduce population growth. Nevertheless this program started in 2004 and is therefore irrelevant, since our data are older. Similarly to Indonesia, almost 90% of population is Muslim, most of the rest being Hindus. Islam is also state religion since 1977 when the country gave up secularism. (Worldmark Encyclopedia of Nations, 2007)

Bangladesh has large potential natural gas reserves and that gained it some interest from foreign investors. Nevertheless about 60% of population still works in agriculture with rice being dominant crop. Although Bangladesh also produces about quarter of the world's supply of jute, only about 20% of the country's GPD is created in agriculture with about half of the Bangladeshi GDP being produced by the service sector. In times our data was collected the country was experiencing steady growth after GDP drop in 1990. This drop was caused mainly by a devastating cyclone that struck Bangladesh that year and domestic political disturbances following first free elections in decades. For children from 6 to 11 years of age school is mandatory in Bangladesh, except for girls living in rural areas who are exempted from this law. Primary education is five years, secondary education is seven years divided into junior, upper and higher secondary education with junior taking 3 years of schooling and upper and higher secondary education taking up 2 years each. There is also Islamic education, although such schools must follow national curricula. Although children under the age of 14 are restricted from working in factories (allowed working in other sectors) this law is not enforced thoroughly and it is not uncommon for children to work even in factories. (Worldmark Encyclopedia of Nations, 2007)

4.1.3 Côte d'Ivoire

The Ivory Coast is situated on the coast of Gulf of Guinea in western part of Africa. It shares boarders with Guinea, Liberia, Mali, Burkina Faso and Ghana. Its area is about twice larger than that of Bangladesh while only inhabiting a little over one tenth of population of Bangladesh – 18 million people were estimated to live in Côte d'Ivoire in 2005. Almost 40% of the population is Muslim. Another quarter is Christian and a little over 10% of people practice traditional indigenous religions. Côte d'Ivoire proclaimed its independence from the French in 1960. Since then it has been led by autocratic president Houphouët-Boigny until his death in 1993. Certainly Houphouët-Boigny was in power at the time data for our base study was collected. (Worldmark Encyclopedia of Nations, 2007)

Key part of the economy of the Ivory Coast is production of coffee, cocoa and tropical woods. This production generates together 40% of the country's GDP. Production of cotton and rubber also plays an important role. Although Côte d'Ivoire experienced massive expansion after gaining independence for the first 15 years, in the 1980s its economy dropped to large extend because of costly debt service. All levels of education are free in Côte d'Ivoire. Primary education takes 6 years, after that the children may study 4 years at lower secondary level followed by 3 years at upper secondary. Children are permitted to work from 14 years of age. Nevertheless this restriction is only enforced in large companies and in civil sector. (Worldmark Encyclopedia of Nations, 2007)

4.1.4 Ghana

Ghana is smaller neighbor of the Ivory Coast situated similarly in West Africa on the coast of Gulf of Guinea. Other than Côte d'Ivoire it shares borders with Togo and Burkina Faso. Population is slightly larger than that of the Ivory Coast with estimated 22 million inhabitants. About 70% of the population is Christian, about 15% Muslim and another 15% of the people of Ghana follow traditional indigenous religions or beliefs. In many areas of the country strong belief in witchcraft is still present. Since the 1960s Ghana was struck by frequent changes of government through various coups both military and civil. The new governments sometimes tried for some kind of democracy, while sometimes they were purely autocratic with no apparent wish for change or claims of democratic elections. The ruler during the time Schultz's (2003) data were collected was chairman of Provisional National Defense Council named Jerry Rawlings. Although at the time he ruled without a constitution he was afterwards elected president in 1993 and reelected in 1996. (Worldmark Encyclopedia of Nations, 2007)

As with the other nations, in Ghana most people work in agriculture as well. It creates about 35% of GDP of Ghana. The key crops are cassava, coco-yams, plantains and yams. Gold is also mined in substantial quantities in Ghana. Ghana is modest oil producer and producer of refined petroleum products. At the time data used by our base study of Schultz (2003) were collected most of economy in Ghana was run by 300 state-owned enterprises. Six years of primary education is both free and compulsory. Primary education is followed by 6 year secondary education. Children are not allowed to work under the age of 15, but many children still work at much younger ages. Daily minimum wage is set at $0.78. (Worldmark Encyclopedia of Nations)

4.1.5 Macroeconomic comparison

Obvious thing that all of our countries have in common is the main criterion we chose them by: the fact that they are all developing countries. Oxford dictionary defines a "developing country" as "A poor agricultural country that is seeking to become more advanced economically and socially."[1] Our countries truly do seek to become more

[1] http://www.oxforddictionaries.com/definition/english/developing-country

advanced economically and socially. They also are mostly agricultural. As described above in all of them majority of population works in agriculture and therefore agriculture plays key role in their economy. They are generally perceived as poor as well and the data prove it.

Table 1: Selected macroeconomic indicators

	Indonesia	Bangladesh	Côte D'ivoire	Ghana
Year of data collection	2007-2008	1996	1985-1989	1985-1989
GDP per capita (current US$)	2,178 (2008)	332 (1996)	833 (1989)	369 (1989)
GDP growth (%)	6% (2008)	5% (1996)	3% (1989)	5% (1989)
poverty at $1.25 headcount	23% (2008)	61% (1996)	14% (1988)	49% (1989)
poverty at $2 headcount	54% (2008)	86% (1996)	35% (1988)	78%(1989)
agriculture in %GDP	14% (2008)	26% (1996)	33% (1989)	49% (1989)
employment in agriculture (%of employed)	40% (2008)	63% (1996)	-	62% (1992)
maternal mortality (per 100ths live births)	250 (2005)	440 (1995)	740 (1990)	760 (1990)
literacy (%)	92% (2008)	35% (1991) /47% (2001)	34% (1988)	-

Source: World DataBank, data retrieved on 3rd May 2014 from http://databank.worldbank.org/
Year of particular indicator record in brackets

Table 1 presents some clear proof our four countries in the sample are all developing countries. The ratio of GDP produced by agriculture is not too large, but the numbers of people employed in agriculture show clear dependence of these countries on their farming industry. Table 1 also points out two representative variables connected to human capital investments. Maternal mortality (here in number of deceased mothers per a hundred thousand live births) is to a large degree clearly dependent on overall state of hygiene which may be connected to education and state of medical care and nutrition. Literacy is than probably the most important indicator of education or lack thereof. From this data it can be seen that our newest data fare the best by far in terms of these two human capital representations, but this large difference is not reflected in poverty numbers to the same proportion.

As Table 1 shows Indonesia seems to be the most developed of our studied countries. It has by far the highest literacy and GDP per capita, lowest margin of agriculture on production of GDP and lowest maternal mortality. Only in terms of poverty does the Ivory Coast fare better. We have to point out that the data from Indonesia are also our newest data, which may be important, and it is true that Indonesia only became richer than Côte d'Ivoire in terms of GDP per capita in the early 1990s (World DataBank, 2014). Still we need to focus on the time our data was collected and in that comparison modern Indonesia fares much better than Côte d'Ivoire at the end of 1980s.

Similarly Bangladesh seems to be the poorest in our sample. It is slightly poorer in terms of GDP per capita than Ghana, but both \$1.25 and \$2 poverty ratios show how many more poor people lived in Bangladesh in the 1990s compared to Ghana at the end of 1980s. Bangladesh fares quite good in terms of maternal mortality compared to the African countries, this may be caused by the fact our data on Bangladesh are from later time than the Ivory Coast and Ghana and medical advancement may have saved many more women. Or possibly the difference may be caused by different cultures between the two regions of the world. Ghana seems to be only slightly richer country (or less poor) compared to Bangladesh in terms of poverty and GDP per capita.

Although Côte d'Ivoire is doing worse than Indonesia by a substantial margin in terms of GDP per capita, its poverty data show there was possibly less inequality between the richest and the most of people at the time of the data collection. All of the countries were going through economic expansion in time of collection of data. Therefore overall economical trend should create no difference.

4.2 Microeconomic data overview

Data we used have been provided by the RAND Corporation. RAND Corporation is a research organization developing solutions for public policy challenges. The data have been available free of charge at the RAND website (rand.org).

Our Indonesian data have been collected between November 2007 and April 2008 and are part of a panel of IFLS – Indonesian Family Life Surveys which is an on-going longitudinal survey. Our data come from its fourth round in particular. It covers 13 out of 27 provinces of Indonesia. For purposes of this thesis it is safe to presume the households questioned were chosen randomly from the population of the 13 provinces.

Our data from Bangladesh are from family and community survey from 1996 entitled Matlab Health and Socio-Economic Survey. It was carried out as the title suggests in Matlab, which is a rural part of Bangladesh, by Demographic

Surveillance System of International Centre for Diarrheal Disease Research, Bangladesh. It is clear that for our comparison of country-specific effect of human capital the fact that in Bangladesh our data are only from rural area poses a problem. We have to adhere to the fact that this data is not only country-specific, but also region-specific in our analysis and it will be mentioned repeatedly throughout the rest of this thesis.

These datasets have been chosen after a long search for possibly useful data. The lack of macroeconomic observations, that is number of countries in our research, is evident, but is given simply by lack of necessary data. Comparison of results across countries is therefore not conclusive, but neither is it impossible or irrelevant. Crucial part of our research is based on models used by Schultz (2003) and in particular his instrumental variables models. To make our comparison with results of aforementioned paper at least a little credible we need as precise compliance between variables used in that paper and in our thesis. The number of instrumental variables used by Schultz (2003) is nevertheless vast and its diversity is extraordinary. As we also present in this chapter we need data on personal measurements (weight, height) as well as on income, but also data on parent education and community data from area of respondent regarding the locality of nearest institutions such as doctor, school or market and prevalence of main diseases in the area. These conditions limit us in our choice of datasets greatly.

4.2.1 Human capital variables

Table 2: Human capital variables overview – mean and standard deviations

Country, Sex, variable	*20-29*	*30-39*	*40-49*	*50-60*	*All*
Indonesia - Males					
Education (years)	9.65	9.28	7.41	6.17	8.41 (4.919)
Migration (1=yes 0=no)	0.387	0.317	0.148	0.093	0.258 (0.438)
BMI	21.09	22.28	22.95	22.31	22.11 (3.670)
Height (m)	1.63	1.62	1.61	1.6	1.62 (0.067)
Inedonesia - Females					
Education (years)	11	8.62	6.16	4.61	7.82 (5.387)
Migration (1=yes 0=no)	0.299	0.197	0.086	0.042	0.165 (0.371)
BMI	22.34	23.99	24.85	24.3	23.86 (4.739)
Height (m)	1.52	1.51	1.51	1.49	1.51 (0.059)
Bangladesh - Males					
Education (years)	1.44	2.16	4.54	3.33	3.23 (4.519)
Migration (1=yes 0=no)	0.222	0.282	0.409	0.34	0.337 (0.473)
BMI	18.85	18.74	18.83	18.69	18.76 (2.193)
Height (m)	1.6	1.61	1.62	1.6	1.61 (0.063)
Bangladesh - Females					
Education (years)	2.89	2.23	1.3	0.59	1.86 (2.967)
Migration (1=yes 0=no)	0.88	0.858	0.87	0.83	0.86 (0.348)
BMI	19.1	19	18.8	17.6	18.76 (2.477)
Height (m)	1.5	1.5	1.49	1.48	1.5 (0.056)

Source of data: RAND Corporation, Obtained on 14th March 2014 from rand.org

Standard Deviations in brackets

Table 2 shows past investment into human capital amongst both sexes in our chosen countries throughout age groups. We can see certain patterns repeating themselves. Although we only include people with incomes in our sample the youngest age group is possibly continuing its education. Except for men of Bangladesh we observe increase in education in recent years making the younger age groups more educated. It is noteworthy that the population of Ghana is more educated, than the population of the Ivory Coast although in our macroeconomical comparison Ghana seems to be poorer. Interesting information is the sharp increase in education of women in all countries. In Schultz's (2003) old datasamples women are substantially less educated than men. This is different though for younger women in the 20-29 age group in both Bangladesh and Indonesia that are actually more educated than their male peers. On average they spend one more year in school compared to male population of those countries.

In Bangladesh more than two times more women migrated from their birthplace than men. This has probably some connection to local culture. Truly the numbers for migration among Bangladeshi women seem to be extremely high compared to other sex-nationality groups. We should mention that migrating from village the person was born in to another village is what is represented in this variable. Therefore a close inter-region migration e.g. to neighboring village is included. Although our migration variable is given by counting at least one migration in lifetime the comparison for the younger age groups do not seem to reflect that and in most countries it is the youngest group that migrate the most. This may be given by the socio-economical changes in previous years in all of these countries, where job opportunities in urban areas became better. The largest difference towards young people migrating may be observed in Indonesia. We might speculate that the governmental program to move people from overpopulated areas to emptier islands that was run in recent years in this country may be more intriguing for younger people since they do not yet have ties to the region of birth.

Body mass index tends to increase with age therefore we cannot tell whether there is any improvement in BMI over time. The differences amongst people of the Ivory Coast and Ghana, from sample used by Schultz (2003) are not too big. Average BMI in Indonesia is slightly higher and average BMI in Ghana somewhat lower. It wouldn't be surprising if higher BMI of Indonesia was caused by better nutrition status of its population and more importantly the low values of BMI in Bangladesh may very probably be caused by malnutrition since it is the poorest country out of the four in question. Notable difference is in trend of BMI of females between Indonesia and Bangladesh. Where in Indonesia younger girls tend to have lower BMI and that grows with age, the opposite is true for women in Bangladesh. Our speculation would be that that is partially due to different social conventions in those countries and partially due to different genetic dispositions.

We can observe some growth in terms of height amongst our age groups. In that younger groups have grown taller than older. Schultz (2003) gives credit for this increase to improved economic situation in the Ivory Coast during the 1970s compared to generation before them and generally better malnutrition situation in Côte d'Ivoire compared to Ghana. We see a similar increase in height between the oldest and second youngest group (20-29) in our data as well and therefore we do not give much weight to these claims.

4.2.2 Other explanatory variables

Other than human capital variables we use other explanatory variables in our models. That is variables that explain income of our sample population, but are not deemed as human capital, e.g. they are not affected by decisions of our sample population. In total it is 48 variables in Indonesia and 2 variables in Bangladesh. In Indonesia these are: 20 variables determining region (province) of residence, 26 ethnicity variables, one binary variable determining whether the person lives in an urban area (=1) or a rural area (=0), and age. Since all of our data from Bangladesh are from a rural region called Matlab we do not use any regional variables, neither do we use a binary variable for urban or rural residence. Also ethnically there are little to no differences in the area, so the questionnaires used for collecting data do not even ask about ethnicity. The only 2 explanatory variables we use for our Bangladeshi sample are therefore age and a binary variable determining whether the religion of respondent is Islam (=1) or other (=0). This variable serves a similar purpose to the ethnicity variables in our Indonesian sample. While religion and ethnicity should obviously not be reflected upon on wage or declining a job applicant and lower wage based on these factors would be viewed as unethical discrimination in developed world, we do leave it in the wage function because we do not know how common such discrimination is in our countries and we cannot rely on it not being a major factor.

There are some differences in the variables used by us and the ones used by Schultz (2003). Since this may affect the final results of our comparison we shall have a look at those and explain the reasons behind them. Schultz (2003) uses similar exogenous explanatory variables such as rural/urban residence binary variable, age and ethnic group (replaced by religion in our Bangladeshi sample), but he adds two more variables: average annual rainfall and a binary variable checking whether the interview occurred during rainy season. Although seasons affect our two sample countries greatly as well as Côte d'Ivoire and Ghana we use income per year instead of last month's wage, which should balance the seasonal differences. Annual rainfall is used in Schultz's (2003) data because his data was collected over period of several years. Since all of our data are collected in one year there is no reason for such variable to be included.

4.2.3 Instrumental variables

Further we use instrumental variables to explain human capital in our IV estimation. These variables do not have a direct connection to income of individuals, but are connected to human capital variables. The instrumental variables used for Indonesia are: 4 local market food prices (particularly rice, fish, oil and sugar), years of education of mother and father, a binary variable being 1 if there is a school in the village or town of residence and 0 otherwise, a binary variable being 1 if there is source of water directly in the place of residence of the respondent and 0 otherwise and two binary variables determining whether parents worked in agriculture for most of their lives (=1) or not (=0). Similar variables are also used for Bangladesh the differences are: the 4 local food prices are price of rice, potatoes, salt and oil, and further we add a binary variable being 1 in case there is a market directly in the village/town and 0 if there isn't. Information regarding availability of market was not included in Indonesian data.

The differences in instrumental variables are somewhat larger than the differences in the exogenous explanatory variables. To preserve as much match between our models we focused on datasets that included at least similar variables. That limited our sources greatly and yet we still did not achieve a 100% match. Since parent education had the most explanatory power in Schultz's (2003) models we only use datasets that also include this variable. Similar can be said about the binary variable checking parents did or did not work in agriculture for most of their lives. Also a variable investigating water and sanitation conditions in the respondents residence is included in both our and Schultz's (2003) models. We had to use smaller number of local food prices than that used by Schultz (2003) (8 and 11 respectively) because our information comes only from families where these products were actually purchased in the last month. Including more food prices would limit our sample size too greatly.

We had to make even larger compromise with the remaining instruments. Schultz uses as instruments distance to Clinic or Hospital, distance to Nurse or Doctor, distance to Permanent Market, distance to Primary school and distance to secondary school. While we did have information about some of these, we did not deem them reliable. Some of these distances were included in our datasets, but only for those institutions that were not located directly in the village/town of residence. They were measured in time and rounded to 15, sometimes 30 minutes. Further there were not large differences among these, with the longest journey taking 90 minutes.

That is why we decided only to use a binary variable reflecting whether there is a primary school (and market in case of Bangladesh) in the village/town. Since vast majority of our sample does have a primary school in town and about one third of our Bangladeshi sample has a market in town, we believe this change should not create a significant difference. The smaller number of instrumental variables though does call for concerns however.

Probably the most problematic instrumental variables were diseases and illnesses in the area. Schultz (2003) uses three binary variables determining whether there was a large occurrence of malaria, diarrhea and measles in vicinity in the last three months before the interview. Since these illnesses are continent specific and are not such a severe threat in Asia (maybe excluding diarrhea) we wanted to include other possibly threatening illnesses. Unfortunately such information was not included in our datasets and therefore we had to completely drop these instruments.

Schultz (2003) also had the option to check for local circumstances from the place of residence at childhood in case of those respondents that have migrated. Since most of the investment into human capital happens in childhood the status of the instrumental variables at that time arguably affects the size of the investment more than current state. Therefore Schultz uses mean value of the community variables from the birthplace of those who migrated instead of the current values. We did not do that change. We could not change the values of the variables because we do not have the information regarding the birthplace for each individual in our samples. Secondly we also did not see this method as very reliable since the birthplace conditions might have changed since the childhood of the person asked in the questionnaire deeming the information completely false and unrelated to the person in question.

4.3 Data issues

We have had to tackle several issues with our data. Some could be solved, while some problems were unsolvable and may affect our results. Firstly our datasets were obtained in bulks of variables and observations have had to be linked between those, we can’t ensure that all the observations have been correctly labeled, since we did not have any direct contact with the source of the data. Similar problem may have happened in case of unintentional error during data collection like typos in data entry.

We have observed some very suspicious outlier observations and we eliminated those by establishing following, in our opinion logical, restrictions: BMI is between 10 and 100; education is positive; prices are positive; we eliminated all observations where prices were smaller than one fifth and greater than five times the mean of the original dataset after applying the positive restriction (these outliers were probably created by wrongly noted volumes of the foods purchased). We believe this restriction is perfectly reasonable and the data prove it by new extremes after applying this restriction becoming a lot closer to the mean that the imposed restriction. We restricted income to be no more than 100 times more than the pre-restriction mean of the sample. We believe all the aforementioned restrictions to be fixing unintentional mistakes during collection of the data as mentioned earlier. We dropped observations showing income below one fiftieth of the value of mean. This restriction nevertheless was not imposed due to fear of error in data entry, but because we want to be, by design, only looking at individuals with actual income, not individuals who were clearly unemployed for majority of the year. In cases where we deemed it negligible we loosened these restrictions. The restriction imposed did not limit number of observations of any of our datasets by more than 5% in total.

5 Empirical results

In this chapter we present the results of our statistical modeling. In some cases we also mention results of Schultz (2003), but complete comparison of the models is included in Chapter seven.

5.1 Correlation among human capital

To satisfy our systematic similarity to the paper of Schultz (2003), that we are following, we shall begin the explanation of our estimation results by noting that different forms of human capital may be correlated with each other. We present these correlations along with correlations with logarithm of income in Table 3.

Table 3: Pairwise correlation of human capital and logarithm of income

Indonesia (3496\2193)					
Male\Female	Log (income)	Education	Migration	BMI	Height
Log (income)		0.3763 (0.0000)	0.0154 (0.4702)	0.0794 (0.0002)	0.1448 (0.0000)
Education	0.2847 (0.0000)		0.2138 (0.0000)	-0.0080 (0.7091)	0.2029 (0.0000)
Migration	-0.0105 (0.5334)	0.1348 (0.0000)		-0.0542 (0.0111)	0.0427 (0.0454)
BMI	0.2116 (0.0000)	0.1431 (0.0000)	- 0.0229 (0.1765)		-0.0537 (0.0120)
Height	0.0819 (0.0000)	0.1911 (0.0000)	0.0227 (0.1803)	-0.0882 (0.0000)	
Bangladesh (1107\1323)					
Male\Female	Log (income)	Education	Migration	BMI	Height
Log (income)		0.1764 (0.0000)	0.0181 (0.5114)	0.1198 (0.0000)	0.0097 (0.9710)
Education	0.3055 (0.0000)		0.0991 (0.0003)	0.1886 (0.0000)	0.1232 (0.0000)
Migration	0.1194 (0.0275)	0.1194 (0.0275)		0.0767 (0.0050)	-0.0114 (0.8643)
BMI	0.2328 (0.0000)	0.2435 (0.0000)	0.0315 (0.5618)		-0.0106 (0.7009)
Height	0.0729 (0.1795)	0.1348 (0.0127)	0.0837 (0.1230)	-0.1079 (0.0001)	

Source of data: RAND Corporation, Obtained on 14th March 2014 from rand.org

No of observations / p-values of correlations in brackets.

The results may appear rather surprising since the correlations in our data do not hold much similarity to correlation of Schultz (2003). Out of the 24 pairwise correlations between human capital variables 14 are positive and statistically significant at the 5% level of significance. The lack of significant correlation between height or BMI and migration is not surprising since there is no reason to see any logical link between these variables. Insignificance of correlation between income and migration in Indonesia can be explained by the governmental program to move people from overpopulated areas that was executed shortly before our data were obtained. If the reason people moved were not for better job, then it is no surprise that there is little or no correlation with income. In case of women in Bangladesh we can only guess that the reason they move is not for job but probably for better marriage options or similar social and cultural reasons, we already mentioned such effect is probable based on the high difference in migration in Bangladesh between men and women (see sub-Chapter 4.2.1.). Probably the biggest surprise is insignificance of correlation between height and BMI amongst Bangladeshi women. BMI should by design be negatively correlated to height.

If we look at the comparison with Schultz (2003), the correlations in Schultz's (2003) data are more significant from statistical perspective, but unfortunately the reasons for correlations between some of the variables are not explained. For example Schultz finds, but does not explain positive statistically significant correlation between height and migration. Also Schultz does not find statistically significant correlation between height and BMI, but he himself mentions that there should be negative correlation by design between those two variables (Schultz, 2003, page 10). Otherwise the sign and to a point also the size of correlations are similar to correlations in our data from Asian countries, with the exceptions noted in the previous paragraph.

5.2 Basic OLS model

Next we present our ordinary least square (OLS) estimates of coefficients of human capital based on linear log income equation. These estimates are presented in Table 4 below. We decided to follow Schultz by allowing for only education in the equation at first and then sequentially adding remaining forms of human capital to show the bias created by omission of the remaining human capital variables. We were sequentially adding variables based on the size of their correlation with logarithm of income.

Table 4: OLS estimates of coefficients of human capital in linear log-income equation.

Country Gender Variable	sequence of equations			
	1	2	3	4
Indonesia - Female (2193)				
Education	0.0997* (17.66)	0.0994* (17.61)	0.1000* (17.65)	0.0969* (17.02)
BMI		0.0094 (1.73)	0.00924 (1.70)	0.0102 (1.88)
Migration			-0.0884* (1.25)	-0.0853 (1.21)
Height				1.769* (4.02)
Indonesia - Male (3496)				
Education	0.0721* (17.75)	0.0672* (16.45)	0.0671* (16.39)	0.0649* (15.75)
BMI		0.0400* (7.78)	0.0400* (7.78)	0.0422* (8.19)
Migration			0.0038 (0.09)	0.0078 (0.18)
Height				1.1607* (4.01)
Bangladesh - Female (1323)				
Education	0.0857* (7.06)	0.0799* (6.53)	0.0805* (6.56)	0.0804* (6.52)
BMI		0.0474* (3.30)	0.0479* (3.33)	0.0480* (3.33)
Migration			-0.0597 (0.60)	-0.0596 (0.60)
Height				0.09485 (0.15)
Bangladesh - Male (1107)				
Education	0.0550* (8.45)	0.0507* (7.72)	0.0477* (7.26)	0.0468* (7.05)
BMI		0.0493* (3.66)	0.0458* (3.42)	0.0478* (3.53)
Migration			0.2435* (3.94)	0.2399* (3.88)
Height				0.4549 (0.98)

Source of data: RAND Corporation, Obtained on 14th March 2014 from rand.org

No. of observations / Absolute value of t-statistics in brackets

*=statistically significant at 5% level of significance

The coefficients for education are as we may expect higher in Indonesia with every extra year of education increasing income of men by 7% and 10% for women. While at the same time the same extra year of education only increases income of people of Bangladesh by 4% for men and 8% for women.

As expected after our findings about correlations of human capital with logarithm of income (see Chapter 5.1) migration is not a significant factor affecting income of both genders in Indonesia and females in Bangladesh. Men in Bangladesh on the other hand gain about 20% extra income if they change place of living. The migration factor for women, although statistically not significantly non-zero, even seems to diminish their incomes. This is especially notable in Indonesia where migrating to a different town/village decreases a woman's income by 14% and this effect is statistically significant at 6% level of significance. This may be happening if women tend to move with their husbands, possibly because of the husbands job or other reasons (like the policy mentioned earlier in Indonesia), and these women may have trouble finding a similarly paid job in the new location.

We see that BMI affects income of women in Indonesia somewhat less than in case of our remaining gender-nationality groups. One standard deviation increases income of female population in Indonesia by about 7%. This affect is even greater for Indonesian men. One standard deviation increases their income by approximately 16%. Slightly smaller increase of income of 12% per one standard deviation of BMI can be observed on our sample of women from Bangladesh. Similar increase in income from one standard deviation of BMI is gained by men in Bangladesh as well. This increase is approximately 13%. That fact is present despite the coefficient of BMI being greater for male Bangladeshi than for Indonesian males, but BMI does not deviate as much in Bangladesh as it does in Indonesia.

Height does not seem to play a significant role for women of Bangladesh from the statistics perspective and also in terms of the size of the coefficient. Nevertheless it does affect income in the remaining gender-country groups. Men in both Indonesia and Bangladesh gain about 1% more money per 1cm of height. Indonesian women on the other hand gain almost 2% more per centimeter of extra height. In terms of standard deviations the difference in height per one standard deviation increases income of female Indonesians by approximately 11%, male population of Indonesia gain about 8.5% more money per one standard deviation of height, and 6% higher income is gained by a Bangladeshi man compared to his one-standard-deviation shorter countryman.

5.2.1 Non-linear equation

There is no substantial reason to assume that the effects of human capital are linear. There may very well be a case of a quadratic effect. Also human capital variables may be mutually complements or substitutes. Therefore we estimate the equation with products of human capital variables pairings and square values of human capital variables. We present results of this estimation in Table 5.

Table 5: Coefficients and t-statistics of non-linear estimation.

Variable	Indonesia Female	Indonesia Male	Banlgadesh Female	Bangladesh Male
Education	-0.114 (0.85)	-0.096 (0.85)	-0.271 (0.73)	-0.155 (0.85)
Education Squared	0.003* (3.46)	-0.001* (-5.95)	0.008* (2.73)	0.003* (2.06)
Education * BMI	-0.001 (0.94)	0.004* (3.18)	0.005 (1.04)	0.005 (1.91)
Education * Migration	-0.004 (0.32)	0.016 (1.70)	0.016 (0.35)	0.005 (0.34)
Education * Height	0.118 (1.35)	0.058 (0.87)	0.111 (0.48)	0.042 (0.38)
BMI	0.356* (2.88)	0.209 (1.74)	-0.437 (1.08)	0.226 (0.59)
BMI Squared	-0.002* (3.85)	-0.002* (2.42)	0.006 (1.72)	-0.004 (1.13)
BMI * Migration	-0.003 (0.21)	-0.016 (1.41)	0.036 (0.76)	0.013 (0.46)
BMI * Height	-0.168* (2.41)	-0.060 (0.99)	0.131 (0.53)	-0.035 (0.19)
Migration	-0.365 (0.20)	-0.796 (0.72)	-4.547 (1.64)	-2.266 (1.29)
Migration * Height	0.255 (0.21)	0.617 (0.94)	2.554 (1.45)	1.388 (1.38)
Height	18.370 (1.46)	1.158 (0.12)	-13.975 (0.63)	11.493 (0.71)
Height Squared	-4.489 (1.18)	0.133 (0.05)	3.097 (0.44)	-3.448 (0.79)

Source of data: RAND Corporation, Obtained on 14th March 2014 from rand.org

Absolute value of t-statistics in brackets

*=statistically significant at 5% level of significance

No. of observations are same as Table 4

The results show in three out of four samples the increasing marginal returns on education. This is presented by the positive coefficient of the quadratic term of education. We were expecting this as multiple previous researchers have found evidence for higher returns from education after primary school (see Psacharopoulos, 1994). Therefore we are even more surprised by the very opposite effect for men in Indonesia. We presume this effect to be strongly offset by the cross effect of education and BMI found in this sample, meaning education is a complement to BMI for this group.

Again as we rather expected the quadratic term of BMI is negative especially in the richer Indonesia. This means there are decreasing marginal returns from body structure. This makes sense since once a person reaches out of the group of malnourished the effects on his work capabilities should diminish substantially. Similarly findings have been found in developed world by Case and Paxson (2008). The only other statistically significant coefficient at 5% level of significance is the cross effect between BMI and height in case of female Indonesians showing BMI and height to be substitutes. Generally however we do not deem this model be very reliable for the high number of statistically insignificant variables used.

5.3 Instrumental variable model with endogenous human capital

5.3.1 Option 1: All human capital variables assumed endogenous

Next we presume that all four human capital variables are endogenous with respect to income. In that case the OLS estimates as described above are biased. We test and assess this assumption of endogeneity in the next Chapter (5.3.2.). To ensure unbiased estimates of coefficients in our model we will use instrumental variables describing local food prices, local infrastructure and parent education and work experience to explain the human capital variables. This method is described as Two-stage least square estimate. The results are depicted in Table 6 below.

Table 6: Coefficients of endogenous human capital variables estimated by instrumental variables

Country - Gender			
Variable	coefficient	t-statistic	p-value
Indonesia - Female (2193)			
Education	0.123	2.21	0.027*
BMI	0.0916	0.7	0.484
Migration	-1.2668	1.37	0.169
Height	8.6634	0.80	0.422
Indonesia - Male (3496)			
Education	0.1795	1.94	0.053
BMI	0.1007	0.39	0.699
Migration	0.0586	0.09	0.931
Height	-13.7290	1.18	0.237
Bangladesh - Female (1323)			
Education	-0.0058	0.11	0.915
BMI	0.3671	3.56	0.000*
Migration	-0.1514	0.20	0.843
Height	7.1440	0.73	0.464
Bangladesh - Male (1107)			
Education	0.0031	0.11	0.911
BMI	0.3307	2.69	0.007*
Migration	0.8332	1.64	0.101
Height	1.5336	0.31	0.759

Source of data: RAND Corporation, Obtained on 14th March 2014 from rand.org

No. of observations in brackets

*=statistically significant at 5% level of significance

The results are not very satisfactory as can be seen mostly by the fact that only 3 coefficients are statistically non-zero at a 5% level of significance. These are female education in Indonesia and BMI in Bangladesh for both men and women. Since these problems were not apparent for the OLS estimates we can explain it by one of two possible reasons: Either the human capital variables are actually not endogenous. And since instrumental variables estimator is by design an inefficient estimator the standard errors become too high. Or our instruments are not explaining the human capital variables properly. The second possibility would be underlined by very low R-squared values obtained during the first-stage equations estimation. Except for education with the remaining human capital variables we could only explain about 10% or less of their variance. In some cases we could only explain 5% of the variance of the variable. On the other hand our instruments were jointly significant at 1% level of significance in all equations. This leads us to believing that we are missing some important instruments to explain the human capital variables properly. We will talk about the quality of our instrumental variables in Chapter six.

To take a quick look at the statistically significant coefficients: One extra year of education increases the income of women in Indonesia by about 12%. That is 2% more than the OLS coefficient told us and that does show a possible bias. One standard deviation in BMI increases income of women in Indonesia by 91% according to our estimate. While at the same time one standard deviation in BMI increases income of a man from Bangladesh by 73%. These numbers would suggest that OLS estimator underestimates the effects of education and greatly underestimates the effects of BMI.¨

5.3.2 Endogeneity test

Now that we have both OLS and IV estimates we can test whether our assumption of endogeneity of human capital variables is correct. We use Haussman-Wu test to do that. See Table 7 for results.

Table 7: Haussman-Wu specification test of exogeneity of human capital variables - test statistics

Country Gender	Education	BMI	Migration	Height
Indonesia Female	3.80* (0.000)	4.09* (0.000)	0.00 (0.998)	2.53* (0.012)
Indonesia Male	7.14* (0.000)	7.69* (0.000)	6.20* (0.000)	3.53* (0.000)
Bangladesh Female	0.98 (0.328)	3.50* (0.000)	0.64 (0.523)	0.16 (0.873)
Bangladesh Male	1.36 (0.174)	3.30* (0.001)	2.77* (0.006)	1.52 (0.130)

Source of data: RAND Corporation, Obtained on 14th March 2014 from rand.org

P-value in brackets

*=statistically significant at 5% level of significance

Our endogeneity test would suggest only BMI to be endogenous in case of Bangladesh. Respectively it also claims migration to be endogenous in case of male population of Bangladesh. At the same time the same test used on our Indonesian data suggests that all human capital variables except for migration (and that is only for women) are endogenous as we suspected originally. In terms of joint endogeneity we can refuse the hypothesis of all of the human capital variables being jointly exogenous at 1% significance level for all four samples.

5.3.3 Option 2: Human capital variables assumed endogenous based on endogeneity test

Similarly to Wu-Haussman test conducted by Schultz, our test also shows endogeneity for only some of the human capital variables in case of Bangladesh. We will therefore repeat our 2SLS estimation assuming only BMI to be endogenous in case of Bangladeshi females and BMI and migration to be endogenous in case of Bangladeshi males. See the new coefficients in Table 8.

Table 8: 2SLS estimation results with some human capital variables assumed endogenous

Country - Gender			
Variable	coefficient	t-statistic	p-value
Bangladesh - Female (1323)			
Education	0.0510	3.15	0.002*
BMI**	0.2893	3.91	0.000*
Migration	-0.1527	1.34	0.179
Height	0.6049	0.86	0.391
Bangladesh - Male (1107)			
Education	0.0155	1.25	0.211
BMI**	0.2988	2.87	0.004*
Migration**	0.8153	1.81	0.071
Height	1.4850	1.84	0.065

Source of data: RAND Corporation, Obtained on 14th March 2014 from rand.org

No. of observations in brackets

*=statistically significant at 5% level of significance

**=variable is assumed to be endogenous

We see slightly lower coefficients of BMI now that the remaining variables are assumed exogenous and estimated as such. And education has become statistically significant for women in improving their income. One extra year of education increases their income by about 5%. In terms of BMI one standard deviation of BMI increases income by 72% for women and 66% for men.

6 Overidentification and instrument strength

In this Chapter we diverge from following Schultz and run further tests to find out whether there is a problem with our instrumental variables estimations. Case of overidentification may happen when one or more of our instruments are correlated with error term in our structural equation. This means it is not a good instrument and should rather be included in the structural equation. In our case it means that some of our instruments affect income directly. The strength of instruments is another thing that may affect our results and we should test it. Weak instruments cause the endogenous variables to be improperly explained and that may easily ruin our models. In our case weak instruments are such instruments that do not have much effect on human capital of our sample population. We add this chapter to make sure that we have followed our base paper of Schultz (2003) properly in terms of correct instrument and that the base paper itself uses good instruments to reach its results. Table 9 summarizes results of our test for overidentification restriction and test for weak instruments

Table 9: Sargan and Anderson test statistics for instruments used

country - gender	Sargan Test	Anderson statistic
Indonesia - Female	15.316 (0.0179)	4.159 (0.7613)
Indonesia - Male	12.118 (0.0594)	2.001 (0.9598)
Bangladesh - Female	40.389 (0.000)	7.537 (0.4799)
Bangladesh - Male	18.157 (0.0113)	12.758 (0.1205)

Source of data: RAND Corporation, Obtained on 14th March 2014 from rand.org

P-values in brackets

6.1 Overidentification tests

Firstly we look at overidentification. As mentioned earlier a possible problem occurs when instrumental variables in system of equations correlate with error term of the structural equation making it not a good instrument, but rather a good either exogenous or endogenous variable included in the structural equation. To test this problem, the model must be overidentified, meaning there are more exogenous variables (instruments) in the model than there are endogenous variables. We used Sargan test statistic to test overidentification restrictions. The null hypothesis of the test is that the instrumental variables are good ones. In other words they do not correlate with the error term in the structural equation. Table 9 shows the test statistic and the p-value for this test. From the results it is clear that we can reject the null hypothesis for 3 out of the 4 samples we have at 2% level of significance and we can reject it even in the last case at 6% level of significance. This clearly shows our instruments are not very good ones because they do not comply with the condition of not being correlated with the error term of the structural equation – they affect the income directly. We could argue for example that the price variables represent local price index too perfectly and that local prices of products in general affect the income (this may easily be the case for farmers). Further we could argue that the population of our sample not having a source of drinking water in the house may suffer with worse hygienic conditions and they get more often rejected at job interviews due to that. All of this is only a speculation though and we leave studying the exact causes of the problem for further research.

6.2 Weak instruments

Another problem considering our instrument may be that the instruments are weak. By weak instruments we mean that they do not explain the endogenous variables we want them to explain enough. Most instrumental variables estimators in reaction to weak instruments do not estimate the coefficients very well. Testing weak instruments is quite complicated in that the volume of necessary explanatory power of instruments may depend on the variables in question. Some tests nevertheless do exist. We chose to use a generally accepted Anderson canonical correlations test. Again the test statistics are shown in Table 9. The null hypothesis of the Anderson test is that the instruments are weak. Again from the p-values observable in Table 9 we cannot reject the null hypothesis at 10% significance level in any of the 4 country-gender specifications. These results suggest that our instrumental variables do not explain the endogenous human capital variables very well. We would need some other, better instrumental variables to truly explain our human capital variables. There is no reason to believe from this test alone that our instruments do not relate to the human capital variables, only that they do not explain them sufficiently. The knowledge that our instrumental variables are weak strongly undermines credibility of our instrumental variable estimations.

7 Comparison with the base study

In this Chapter we will compare the results of our estimations based on data from Bangladesh and Indonesia with the results described by Schultz (2003) based on data from Ghana and The Ivory Coast. We described the main discrepancies in the data from the countries as well as the social, political and economic differences between the countries at the time of the data collection in Chapter 4. The differences in the results may often come from these differences in country situation and the differences in datasets as we will try to explain in this Chapter.

7.1 Pairwise correlations

Firstly we present a small comparison of the correlations found between the human capital variables and logarithm of income, or wage in case of Schultz (2003). Firstly Schultz (2003) finds significant (at 1% level of significance) positive correlation between logarithm of wage and all four kinds of human capital. Our results on the other hand only do not show significant correlation of migration with logarithm of income at 5% level of significance. Only case where correlation of migration and income is significant at the 5% level is in case of men in Bangladesh. In Indonesia this lack of significance can be explained by the governmental program to move people from overpopulated areas to less populated islands in the country. This program may have been enough of an incentive for people to move and the effects of better job opportunities for migrating was diminished so much that correlation disappeared. In case of women in Bangladesh we found an incredibly high rate of migration, which we assume is given by social conventions in Bangladesh, perhaps moving for marriage. Further we did not find significant correlation between logarithm of income and height in Bangladesh for both genders. Economic explanation of correlation between income and height is complicated, although frequently observed. Schultz (2003) actually does not present explanation for finding this correlation. Our speculation would be that since we observe a very rural region of Bangladesh we have a lot higher number of farmers in our sample compared to our Indonesian data. If than those farmers own their own farms they do not need any psychological power status towards their employers. And height may have such psychological effect that grant the taller people better job offers from employers, but do not give them actual advantage when self-employed during highly physical work on a farm. Nevertheless please bear in mind this is pure speculation and we do not hold or know of any good proof for any of the previous claims.

The correlation coefficients between logarithm of wage and education are similar. The coefficients are between 0.21 and 0.48 in Schultz's data, while the same coefficients are between 0.18 and 0.38 in our datasets. The correlation we found between migration and logarithm of income for male Bangladeshi is also within boundaries of what was found by Schultz. We find somewhat lower correlation between BMI and log (income) in both countries – 0.08 and 0.12 respectively while Schultz finds 0.21 and 0.27. And we find similar correlation of BMI and income for men. Also correlation of height and logarithm of income in Indonesia is not much different from findings from both countries studied by Schultz.

Further we find significant positive correlation between migration and education in all country-gender groups. Similarly Schultz finds correlation between migration and education in 3 out of 4 of his datasets. He does not find the correlation for women in the Ivory Coast significant at 5% level of significance, but he would find it significant at the 10% level. The coefficients also do not differ very much from what we found in our data.

BMI has significant positive correlation with education at 5% level of significance for all four data groups used by Schulz. We nevertheless do not find correlation between education and BMI significant for our female sample from Indonesia. There is no realistic explanation we can give for this exception. The coefficients of the remaining of these correlations are similar between our and Schultz's data.

Schultz further finds positive correlations between BMI and migration and between height and migration in all four of his datasets to be significant at 5% level of significance. We only find significant correlation between BMI and migration at 5% level of significance in our female Bangladeshi dataset and just significant correlation between height and migration at 5% level of significance in female dataset from Indonesia. We do not see any reason for correlation between variables explaining physical measurements and variable saying whether a person has moved from birthplace or not. The only reasonable explanation we see could be strong correlation between both of the variables and some third variable – in this case for example education.

Schultz (2003, p. 10) admits that BMI is structured so that there should be slight negative correlation between BMI and height. We find this negative correlation statistically significant at 5% level of significance in all of our datasets but female Bangladeshi. Schultz on the other hand does not find this significant correlation. In his data from Ghana for both genders the correlation between BMI and height is even positive and significant at 5% level of significance. This finding raises questions regarding the specifics of the datasets.

7.2 OLS estimation

Next we compare the results of our OLS (Ordinary least square) estimator in approximating the log-income equation as shown in Chapter 3 with a similar estimation by Schultz. Firstly in terms of significance of coefficients of human capital Schultz does find all of the coefficients statistically significant at 5% level of significance, with the exception of height in his sample of women from the Ivory Coast. We on the other hand have found several human capital variables not statistically significant at 5% level of significance. It is Migration in all but male Bangladeshi sample. Again we presume the reasons to be as we explained them earlier in Chapter 7.1. The migration of people in Indonesia was probably in a large part not for better job opportunities and therefore better income but rather the Indonesians migrated due to governmental program to move people from overpopulated areas. And the migration of women from Bangladesh is so common that we presume it is again not due to better income opportunities but presumably due to social and cultural reasons. Not only was money not the reason for moving, but more importantly based on these results it looks like this migration does not yield any extra revenue.

Furthermore we don’t find significant effects of BMI on income for women in Indonesia at the 5% level of significance. It would however be significant at 6% level of significance. Nevertheless even if it would be significantly different from zero from statistics perspective, the coefficient is still substantially lower than in case of other country-gender groups. We could explain this by saying that Indonesian women as women from a more developed country than in our other samples do not work in so many manual jobs and therefore their physical fitness does not play a major role in their income. This is slightly backed by somewhat larger effect of height on their incomes than for other groups since height possibly increases income due to psychological effects that may be larger outside of such manual jobs.

Last two coefficients where we did not find statistical significance at 5% level of significance is height in Bangladesh for both men and women. Again we explain this by presuming that the main effect of height is psychological – that it gives taller people a psychological advantage in diplomatic negotiations over wage. Since our data are from a rural part of Bangladesh it's safe to presume a very large portion of our sample consists of farmers. For farmers these negotiations are much simpler, in case of farmers working on their own farm there are almost none. Therefore height does not increase their income dramatically.

The coefficients of education are clearly lowest in the Ghanaian sample, where one extra year of education creates almost 4% increase in wage. The situation is quite similar for men in Bangladesh where each extra year of education brings over 4.5% more income. In all other cases however the coefficients are clearly larger. Every additional year in schools gives between 6.5% and 11% more income. These results point at more jobs being created outside Ghana in fields where education is necessary, where Bangladesh even as the poorest country in our sample creates still enough of these job opportunities that women living in this country gain 8% extra income with every year of education. Clearly the demand for education is greater in this country amongst working women than amongst working men. Interestingly larger coefficient for education for women than for men can also be observed in our Indonesian sample. Schultz (2003) reports opposite results – men in his samples according to his results gain more from every extra year of education than women.

Apart from already mentioned insignificance of BMI in relation to income for women in Indonesia, the effects of BMI in all of our other samples are similar – between 4% and 5% increase in income per extra point of BMI. Similar effects were found by Schultz for men in the Ivory Coast and women in Ghana. Slightly larger effects of 5.3% increase in wage per point of BMI were found in his male sample from Ghana and even larger increase of 6% per point of BMI was observed in sample of women from the Ivory Coast. Since women in the Ivory Coast also have one of the largest standard deviations in BMI of our samples, the differences between particular women appear to be highest there with about 27% increase in wage per one standard deviation of BMI. In all the other samples the increase is between 10% and 20% of extra income per one extra standard deviation in BMI. Overall the differences in effects of BMI between countries do not appear to be very large with the exception of women in Indonesia where BMI has very small or no effect on income.

Migration as mentioned above is not statistically significant at 5% level of significance in 3 out of 4 of our samples. The effects even in Schultz's results do not appear very consistent between samples. Migrating brings 71% extra wage for men in the Ivory Coast and 89% for women while it gives only 35% increase in wage to men from Ghana and 53% increase in wage for women in the same country. We found the effect of migration to increase income of Bangladeshi men by about 24%. We should note that the standard deviation of migration is fairly similar in all of these samples. Migration clearly appears to be very specific variable in that reasons for investing into migrating elsewhere are vast and not only monetary. The cultural differences affect this variable so much that in our opinion there is no reason to look for any clear and obvious trends in size of its effect on income.

Height appears to also be a problematic variable to interpret. We did not find its effects on income statistically significant at 5% level of significance for both of our data samples from Bangladesh and Schultz (2003) did not find it statistically significant for women in the Ivory Coast. The coefficients are also not very similar across countries. One centimeter of height increases wage of men in the Ivory Coast by 0.9%, while the same centimeter increases wage in Ghana by 1.5% for men and 1.3% for women. In Indonesia the effects are even larger with one centimeter increasing income by 1.2% for men and 1.8% for women. Again the standard deviations are similar across all samples being between 0.059 and 0.067. There does not appear to be any lead on interpreting the differences in effects of height.

Overall our OLS estimates for Bangladesh and Indonesia do not differ too much from the estimates of Schultz (2003) for the Ivory Coast and Ghana in terms of BMI effects on income. Effects of education seem somewhat larger in more developed countries, while migration and height effects on income seem to be very country-specific. We have to note however that any comparison based on a very limited sample of only 4 countries, or 8 country-gender groups are completely insufficient for any conclusion regarding trends in human capital effects.

7.2.1 Non-linear model

The results of both our and Schultz's expanded log-income equation covering also quadratic terms of human capital variables and cross effects are not showing too much. Both we and Schultz find proof for increasing marginal returns from education which was expected for it has been previously established that returns from education increase after primary school (Psacharopoulos, 1994). The only negative coefficient for the quadratic effect of education suggesting the opposite is the case of men in Indonesia. In this sample however we also found a strong complementary effect of education together with BMI. This means physical stature has to go along with education in this country to be effective. In three out of four datasets from Ghana and the Ivory Coast the opposite effect – a substitution one has been found between BMI and education.

We have also found a statistically significant (at 5% level of significance) negative coefficient of the quadratic term of BMI in Indonesia. This suggests that BMI has decreasing marginal returns. In no other country is such an effect observable however leading us to believe that only Indonesia is a rich enough country to actually have a substantial part of population out of the hunger risk and in a good physical shape which grants them enough work efficiency. Other than this both our and Schultz's results are struck by insignificance of most of the coefficients at the 5% level of significance not making space for any further complementary or substitutions claims neither showing and trend in marginal returns on human capital variables.

7.3 Instrumental variables estimates

Next we would like to compare results of similar estimates to previous sub-chapter, this time estimated by using Two-Stage Least Square estimator (2SLS). OLS is unbiased under the assumption, that all right-hand side variables of the equation are exogenous. Once we assume some of them are endogenous, OLS is no longer unbiased. The assumption that human capital variables are endogenous seems logical. For example better nutrition may lead to better physical stature, which may lead to higher output at work, which according to economic theory should lead to higher incomes. But on the other hand higher income leads to better nutrition. Similar arguments also apply to the remaining human capital variables used in our research.

Endogeneity of a variable may be tested. We used Hausman-Wu test, similarly to Schultz (2003) to test endogeneity of our human capital variables. Schultz (2003) only finds BMI and height in Ghana for both genders and height for women from the Ivory Coast and BMI for men from the Ivory Coast to be endogenous. However we found endogeneity for all human capital variables in case of men in Indonesia. Next we found education, BMI and height to be endogenous in case of women in Indonesia. For Bangladeshi males BMI and migration were endogenous. And lastly in case of women from Bangladesh, only BMI appeared to be endogenous. All of these at 5% level of significance. In total then BMI was deemed endogenous by the testing in all cases except for women from the Ivory Coast and height was endogenous in all of our country-gender pairings except for men from the Ivory Coast and both women and men from Bangladesh. Education was only endogenous in Indonesia and migration was only endogenous for the two male samples tested by us. This shows already, that there are some major differences in the data between our samples and data used by Schultz (2003).

However since we would from theory presume all of the human capital variables to be endogenous we shall compare our 2SLS estimates with the 2SLS estimates made by Schultz (2003) under the assumption that all four human capital variables are endogenous. Firstly Schultz (2003) finds all the effects of human capital, except for two, at least borderline statistically significant at 5% level of significance. The two effects being not significant at the 5% level are height effects in both men and women sample from the Ivory Coast. In our estimations however we only found three out of the 16 coefficients to be statistically significant at the 5% level of significance. They are education in estimation for female Indonesians and BMI in both male and female samples from Bangladesh. This along with the confirmation of endogeneity of human capital as discussed above gives clues to flawed methodology or data used somewhere along the estimation.

The expected income gain from one year of education of women in Indonesia is, according to the 2SLS estimation, 12%. This is almost 3% more than the OLS estimation showed. Since Schultz does not find education to be endogenous by the Hausman test it is no surprise that the coefficients he gets from the 2SLS are fairly similar to OLS and also along the lines with our OLS estimation of this effect.

The gain from 1 point of BMI in Bangladesh is approximately 3.5% for both men and women, which is about 1.5% smaller effect than estimated by OLS. The estimates obtained by Schultz however shows opposite bias of OLS. His 2SLS coefficients for BMI are larger than those estimated by OLS in all 4 cases and it shows about 8-16% of wage increase per 1 extra point of BMI.

To keep our similarity with the paper Schultz (2003) we also tried estimating another 2SLS model with assuming that only some of the human capital variables are endogenous as was show by the Wu-Husman test. We only estimated this model for our Bangladeshi samples since in case of Indonesia all human capital were deemed endogenous by the test (except for migration for women). The results however did not change very much from the case where all human capital variables were assumed endogenous. The only new coefficient that was also statistically significant at the 5% level of significance was education in the female sample. Its coefficient showed that one extra year of education grant the Bangladeshi women a 5% increase in income. Note that education is assumed exogenous in this case. All the other coefficients of human capital variables, both exogenous and endogenous were not statistically significant at 5% level of significance. BMI kept being statistically significant as in the previous case and its coefficient slightly decreased. Results of Schultz (2003) do not show any large change from the model with all human capital variables assumed endogenous in any of the coefficients or their statistical significance.

Since the results of our 2SLS estimation was so different from results by Schultz (2003) we took a closer look at the instrumental variables used in our model. We tested their strength and overidentification restriction. This testing was not performed by Schultz (2003) and according to our results wrongly so. The testing showed strongly that the instrumental variables used are far from ideal having problem complying with both of necessary properties of instrumental variables. This shows that the problem is either that the instrumental variables of Schultz were not very good or at least need to be tested themselves, or there was a problem transitioning the use of same instrumental variables in our research.

8 Conclusion

We have conducted statistical testing of effects of human capital on income in developing countries. We followed paper of Schultz (2003) using new data to find further clues to reasons behind differences in size of effects of human capital in various countries. We estimated the effects of four variables representing human capital on income. The human capital variables were: education representing number of years spent in school, migration representing moving to a different place of residence possibly for the reasons of better work opportunities, BMI representing investment into nutrition and short-term health and height representing investment into child nutrition and long-term health. We conducted an estimation of income equation using OLS where we assumed all of the human capital variables to be exogenous. We also estimated the effects of human capital on income using 2SLS estimator when we assumed human capital variables to be endogenous. In this case we explained the human capital variables by instrumental variables that were as similar to the ones used by Schultz (2003) as possible for validity of comparison of results.

Assuming human capital variables to be exogenous we arrived at similar results as those by Schultz (2003). In other words: our OLS estimation coefficients were in accordance with Schultz (2003). Education seemingly was more valuable in countries with higher GDP and lower poverty. Effects of BMI on income were similar in all of the estimations. Effects of height and migration however were very different and seemingly random. We note that the very limited sample of 4 countries do not grant us any base for claims of any trends or lack of them in effects of human capital on income.

We also estimated an income function including quadratic and cross effects of human capital variables. Over both our and Schultz's (2003) results they showed an increasing marginal returns on education in 7 out of 8 samples. Also decreasing marginal returns on BMI has been found only in Indonesia – the richest country of the four we are comparing.

However when we assumed human capital variables to be endogenous and employed 2SLS to estimate the model the results were very different from those presented in the paper of Schultz (2003). Most coefficients were not statistically significant. The remaining coefficients were not in line with the results of Schultz (2003).

The endogeneity of human capital variables were tested in both our thesis and the paper of Schultz (2003) by Hausman-Wu endogeneity test. While Schultz does not find endogeneity of education and migration, we find all four human capital variables to be endogenous in our Indonesian male sample and three out of four in Indonesian female sample. In our male Bangladeshi sample BMI was deemed endogenous by the test and BMI and migration were accessed as endogenous for women in Bangladesh. We also, similarly to Schultz (2003), conducted separate estimation only assuming endogeneity for variables that were deemed endogenous by the Hausman-Wu test. The results however did not change drastically from the model where all human capital variables were assumed endogenous.

The reasons behind so large difference in results of the 2SLS estimation were uncovered by testing the instrumental variables. When testing their strength we used the Anderson test to access the explanatory power of the instrumental variables on the human capital variables. In all four cases we could not reject the hypothesis that the instruments are week even at 10% level of significance. That is the instrumental variables used did not fulfill the function of explaining the human capital variables very well and we would need to use some different instrumental variables. Secondly we tested our instruments for overidentification restriction. That is that the instruments we used are in fact correlated with error term in our first equation explaining the logarithm of income. We used Sargan test to find out whether we faced this problem and in all four cases we rejected the hypothesis that the instruments are OK at 6% level of significance. Our instruments therefore are not very good for they directly affect the logarithm of income in some way.

Similar testing was not performed by Schultz (2003). The results of the tests we performed clearly show that the instrumental variables used were not good. The possible reasons for the large differences in results of aforementioned paper and of our estimation therefore are: Either the instrumental variables chosen by Schultz are wrong from the beginning and different IVs should be used. Or there are specifics of Asian countries in our sample that make these variables bad. Or since we used slightly different instrumental variables from the ones used by Schultz, the change had larger effect on overall performance of these variables than we expected. This could have happened at two instances: We did not use data from the place of origin for people who migrated and since a large proportion of our samples migrated, that may have had a more substantial impact than anticipated. We also did not include any

IVs indicating dangerous illnesses and diseases afflicting the community and our instrumental variables describing institutions in the neighborhood were more simplistic than the ones used by Schultz (2003). The last possible problem of our estimations may be the fundamental validity of the underlying data obtained from RAND Corporation. Since we found very significant outliers in some of the IVs, especially the local food prices, we cannot be sure about the overall quality of the sample, even after we excluded these outliers.

Overall we can say that we found similar results to the ones from the paper by Schultz (2003) in our OLS estimation. Our 2SLS estimation however casts doubts regarding instrumental variables used by both us and Schultz (2003). Since Schultz (2003) does not test the quality of the instruments we cannot say with certainty that the results of this paper are valid.

We did not find any clues of trends in size of effects of human capital on wage with any reasonable certainty, neither did we show otherwise. That remains for further research of the topic. The instrumental variables used in this thesis and in Schultz (2003) may also be tested in future for further verification of our conclusions. Also the correct instruments for establishing a robust model describing human capital effects on income across a population in various world countries, whether developing or developed, remain to be found.

Bibliography

Baum, C. L. & W.F. Ford, 2004. The wage effects of obesity: a longitudinal study. Health Economics, Vol. 13, pp. 885–899

Case, A. & C. Paxson, 2008. Height, Health, and Cognitive Function at Older Ages. The American Economic Review, Vol. 98, No. 2, pp. 463-467

Deolalikar, A.B., 1988. Nutrition and Labor Producvivity in Agriculture: Estimates for Rural South India. The Review of Economics and Statistics, Vol. 70, No. 3, pp. 406-413

Griliches, Z., 1977. Estimating the Returns to Schooling: Some Econometric Problems. Econometrica, Vol. 45, No. 1, pp. 1-22

Haddad, L. J. & H.E. Bouis, 1991. THE IMPACT OF NUTRITIONAL STATUS ON AGRICULTURAL PRODUCTIVITY: WAGE EVIDENCE FROM THE PHILIPPINES. Oxford Bulletin of Economics and Statistics, Vol. 53, Issue 1,pp. 45–68

Kedir, A. M., 2013. Schooling, BMI, Height and Wages: Panel Evidence on Men and Women. Economic Issues, Vol.18, No.2, pp.1-18

Psacharopoulos, G., 1994. Returns to investment in education: A global update. World Development, Vol. 22, Issue 9, pp. 1325-1343

Schultz, T. P., 1982. Lifetime Migration within Educational Strata in Venezuela: Estimates of a Logistic Model. Economic Development and Cultural Change, Vol. 30, No. 3, pp. 559-593

Schultz, T.P., 1997. Wage and labor supply effects of illness in Côte d'Ivoire and Ghana: instrumental variable estimates for days disabled. Journal of Development Economics, Vol. 53, Issue 2, pp. 251-286

Schultz, T.P., 2003. Wage Rentals for Reproducible Human Capital: Evidence from Ghana and the Ivory Coast. Economics & Human Biology, Vol. 1, Issue 3, pp. 331-366

Steckel, R.H., 1995. Stature and the Standard of Living. Journal of Economic Literature, Vol. 33, No. 4, pp. 1903-1940

Strauss, J., 1986. Does Better Nutrition Raise Farm Productivity?. Journal of Political Economy, Vol. 94, No. 2, pp. 297-320

Strauss, J., Thomas, D., 1995. Chapter 34 Human resources: Empirical modeling of household and family decisions. Handbook of Development Economics, Vol. 3, Part A, pp. 1883-2023

Thomas, D. & J. Strauss, 1997. Health and wages: Evidence on men and women in urban Brazil. Journal of Economics, Vol. 77, Issue 1, pp. 159-185

Willis, R. J. & S. Rosen, 1979. Education and Self-Selection. Journal of Political Economy, Vol. 87, No. 5, Part 2, pp. S7-S36

Online publications

Thomas, D. & E. Frankenberg, 2002. Health, nutrition and prosperity: a microeconomic perspective. Bull World Health Organ [online], Vol. 80, No. 2 [cited 2.1.2015], pp. 106-113 . Available from: http://www.scielosp.org/scielo.php?script=sci_arttext&pid=S0042-96862002000200005&lng=en&nrm=iso

Thomas, D. et al., 2006. Causal Effect of Health on Labor Market Outcomes: Experimental Evidence. California Center for Population Research. UCLA: California Center for Population Research. [cited 2.1.2015] Available from: http://escholarship.org/uc/item/0g28k77w

"Bangladesh." Worldmark Encyclopedia of Nations. 2007. [cited 2.1.2015] Available from Encyclopedia.com: http://www.encyclopedia.com/doc/1G2-2586700190.html

"Côte D'ivoire." Worldmark Encyclopedia of Nations. 2007. [cited 2.1.2015] Available from Encyclopedia.com: http://www.encyclopedia.com/doc/1G2-2586700092.html

"Ghana." Worldmark Encyclopedia of Nations. 2007. [cited 2.1.2015] Available from Encyclopedia.com: http://www.encyclopedia.com/doc/1G2-2586700101.html

"Indonesia." Worldmark Encyclopedia of Nations. 2007. [cited 2.1.2015] Available from Encyclopedia.com: http://www.encyclopedia.com/doc/1G2-2586700200.html

Data Source

World DataBank, 2014. Obtained 3.5.2014. Available at: http://databank.worldbank.org

The Indonesia Family Life Survey, 2008. RAND Corporation. Obtained 15.3.2014. Available at: http://www.rand.org/labor/FLS/IFLS

Matlab Health and Socio-Economic Survey, 1996. RAND Corporation. Obtained 15.3.2014. Available at: http://www.rand.org/labor/FLS/MHSS