



LUNDS UNIVERSITET  
Ekonomihögskolan

# Inequality of Opportunity

Determinants of Access to Education

– With a Special Focus on Income Inequality

Author: Johanna Olsson

Supervisor: Therese Nilsson

Institution of Economics

2011-05-25

## **ABSTRACT**

In this study the concept of inequality of opportunity is applied to access to education and investigation is made over what determines this. A particular focus is set on income inequality and the role it plays for access to education. Several hypotheses regarding the relationship between income inequality and inequality in education, used as a proxy for access to education, can be thought of. Four hypotheses are presented in this study however there is no consensus regarding the sign of the investigated relationship.

This relation is empirically investigated in a cross-section analysis covering 64 countries. This study applies a model in which the change is used in both the dependent and the independent variable. The changes are derived over the period 1975-1995 for income inequality and 1980-2000 for education inequality. Hopefully, as this method mitigates the otherwise not uncommon problem of endogeneity, the result will to a larger extent be able to give indication of a casual relationship.

Ultimately this study shows that in a simple model-specification a positive relation between income inequality and education inequality is found. However, as the model is extended to include control variables both the significance and magnitude of the relationship decreases. The control variables included in the model, initial GDP per capital level, initial education inequality, urbanization and the change in average years of education are however statistical significant.

***Keywords: access to education, inequality in education, inequality in income, inequality of opportunity, empirical analysis.***

## **ACKNOWLEDGMENT**

I would like to give a special thank to my supervisor, Therese Nilsson, for her inspiration, support and always valuable advices throughout the whole process of this study. It has all been of great value and importance to me.

## TABLE OF CONTENTS

1. INTRODUCTION.....	3
2. BACKGROUND.....	7
2.1 Access to education .....	7
2.1 Inequality of opportunity .....	8
2.3 Inequality in education .....	9
2.4 Inequality in income .....	11
3. HYPOTHESES .....	12
3.1 Development and democracy level .....	14
4. PREVIOUS RESEARCH .....	15
5. EMPIRICAL SPECIFICATION.....	17
5.1 Dependent and independent variables .....	17
5.2 Control variables .....	19
5.3 Data .....	20
6. EMPIRICAL ANALYSIS.....	22
6.1 Descriptive analysis.....	22
6.2 Regression analysis .....	26
6.3 Sensitivity analysis.....	28
7. CONCLUSION .....	32
REFERENCES.....	34
APPENDIX .....	37

# LIST OF FIGURES AND TABLES

## FIGURES

<i>Figure 1</i> .....	5
<i>Figure 2</i> .....	8
<i>Figure 3</i> .....	9
<i>Figure 4</i> .....	19
<i>Figure 5</i> .....	23
<i>Figure 6</i> .....	24
<i>Figure 7</i> .....	24

## TABLES

<i>Table 1</i> .....	6
<i>Table 2</i> .....	13
<i>Table 3</i> .....	20
<i>Table 4</i> .....	22
<i>Table 5</i> .....	25
<i>Table 6</i> .....	28
<i>Table 7</i> .....	29
<i>Table 8</i> .....	30

## 1. INTRODUCTION

For some decades the issue of inequality has received much attention in economic research. The World Development Report from 2006 covering the subject of Equity and Development states two fundamental reasons why inequality is an important issue to deal with. The first reason lies in the nature of human beings, namely the fact that many of the inequalities that determine an individual's opportunities to shape its own outcome of life are pre-determined and not controlled by the individual itself. That contradicts the concept of justice and fairness. The second reason is that inequalities affect the well-being of human beings. It directly affects the welfare of people who receive less of the distribution in the society but it also affects the entire society as there is evidence that inequality affects growth and development in a negative way.

Inequality has often been measured in outcomes such as income, health, and physical capital, but lately the concept of inequality of opportunity has received much attention in the inequality debate. This concept is closely connected to the first reason outlined above to why inequality is an important subject. Inequality of opportunity relates to concepts such that individuals should be given a "fair chance" or "levelling the playing field".

This study investigates what determines inequality of opportunity, which in this study is applied to education. Several aspects of inequality in education can be thought of. Aletta Grisay states five principles of inequalities in education: *the natural equality principle, the equality of access/opportunity principle, the equality of treatment principle, the equality of achievements principle, and the equality of social fulfilment principle*. The first principle takes the "libertarian" stance and is not concerned with equality, it is argued that liberty suffers from forced intervention. The second principle assumes that the level of education an individual hopes to achieve is determined by its potential or natural talent and it refers to the situation in which access to and success in education is not affected by the individual's socio-cultural background. The equality of treatment principle assumes that all individuals have the possibility to benefit from basic education and it refers and opposes to situations in which unequal quality of education is responsible for unequal achievements. The equality of achievements principle assumes individual characteristics to be modifiable and refers and opposes to situations in where unequal education amplifies the inequalities present at the start. The last principle refers to situations in which individuals leaving the educational system have the same opportunities to use their acquired skills. (European Group for Research on Equity in Educational Systems, 2005, pp. 14-15)

Since this study investigates inequality of opportunity applied to education, the most suitable principle to employ is the equality of access/opportunity principle. Both inequality of opportunity and the equality of access/opportunity oppose to situations where pre-determined circumstances of an individual shape the outcomes of the individual's life. This study uses the Gini coefficient over educational attainment as a proxy for access to education.

New growth theory stresses the importance of aggregated human capital accumulation for an economy's growth. There are several explanations to a prevailing inefficient low level of human capital in an economy. The most common explanation is individuals being credit constrained via imperfect credit markets. The theoretical idea is that poor people do not have the necessary capital and will not either be able to lend via capital markets to fund investments in education. Therefore poor people are constrained in their education investments, and might not be able to invest in education at all or only to a small amount. Hence the more unequal the society is, for a given GDP level, more people are limited in their education investment by the credit constraint and the average level of human capital in the economy is low. (Mejia and St-Pierre, 2008, pp.396-397)

Mejia and St-Pierre suggest another explanation to the relationship mentioned above. In their work they put forward a model based on the idea of differences in the rate of return to human capital investments. The rate of return is dependent upon differences in an individual's opportunity set and specified as an individual's endowment of complementary factors to time and effort in the schooling process. Many of these complementary factors are pre-determined and not controlled by the individual itself, e.g. race, composition of household, parents' education. In their model a negative relationship between inequality of complementary factors and the average human capital in a society is found. (Mejia and St-Pierre, 2008, pp. 396-407) Both these explanations can be drawn back to the concept of inequality of opportunity, as they refer to situations in which an individual's opportunities to shape its outcomes in life, for example acquiring an education, are affected by pre-determined circumstances.

However, empirical analyses of the importance of human capital for growth have not been as conclusive as theory predicts. One explanation to this is put forward by Thomas, Wang and López (1998). Their explanation lies in the fact that research only to a small extent has considered the distribution of education, which theory stresses to be important. As ability differs among individuals and can be assumed to be normally distributed, a skewed distribution of income, education or health will lead to underutilization of people's capacities and hence welfare losses. Thomas and Wang argue that production in

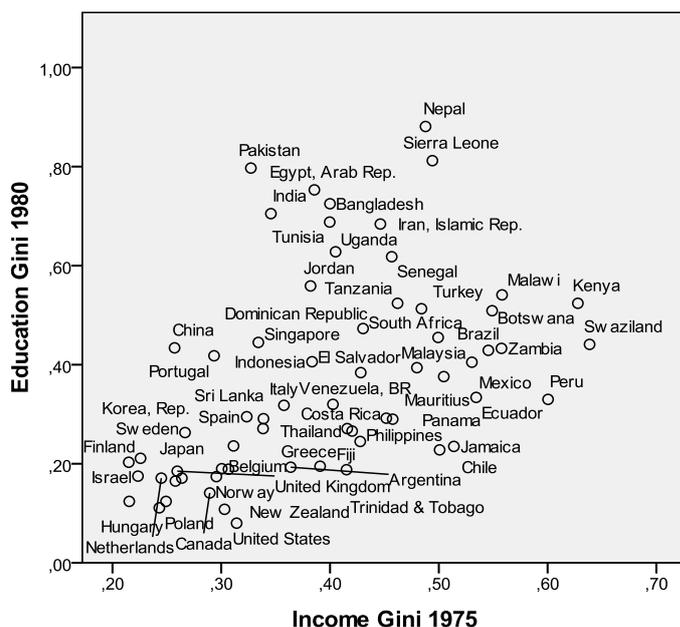
an economy reaches an optimum when the distribution of education coincides with the distribution of ability. Hence inequality of opportunity will make people not fully capable to use their capacity which leads to inefficiency in the economy. (Thomas and Wang, 2009, pp. 36-41)

Empirically, few studies have been done regarding the importance of the distribution of education. One study is carried out by Thomas, Wang and López (1998), who develop a model where both aggregate level and the distribution of education enter the production function. Their empirical study shows that higher inequality in education has a negative impact on per capita income in most countries. Another study made on the subject, by Castelló and Doménech (2001), also shows negative effects of inequality in education on growth rates. Some evidence of the importance of the distribution of education exists and it is therefore important to get better knowledge of what affects access to- and also inequality in education.

There are several factors, identified by numbers of researchers, that can lead to educational inequality; culture, gender, disabilities, globalization, HIV/AIDS, natural disasters, neoliberalism, political economy, politics, poverty, privatization, race or ethnicity, religion, social class, societal values and norms, socioeconomic status and war. (Holsinger and Jacob, 2008, p. 5) However, little attention has been placed on income inequality when studying inequality in education. This study investigates further which variables can be found to determine access to education and sets a particular focus on income inequality and the impact it has on access to education.

The purpose of this study is to analyse the relationship between the income distribution and access to education. Several hypotheses regarding the association between education inequality and income inequality can be proposed. However, there is no consensus on the direction of the relationship. This gives give good reason and an interesting case to investigate this relation empirically. Observations from 64 countries in a cross-section dataset are used in this study, to investigate the issue further.

Figure 1. The cross-country correlation between income inequality 1975 and education inequality 1980.



As can be seen from the cross-country correlation between education inequality and income inequality presented in *Figure 1*, a positive relation between the variables is evident. However, such description of data only gives a first indication of a relation between income inequality and education inequality and several other aspect need to be taken to consideration. First, other potentially important determinants of access to education are not taken into account. Another shortage is the possibility that the relation shown, reflects a relation between an omitted variable that affects access to education that in turn is affected by income inequality, instead of a direct relation between the two variables shown. Moreover, reverse causality is an important issue not taken into consideration in a correlation. This calls for a more rigorous investigation of the relationship.

This study develops a model that further investigates this relation while taking into account the problems outlined above. Investigation whether the result is robust across high and low income countries as well as across democracy and non-democracy countries will also be carried out. Various additional sensitivity analyses are also performed to test the robustness of the result.

The analysis shows that a simple model yields a positive, significant relation between income inequality and education inequality. However, it will be shown that this relation is not robust. Other variables such as GDP per capita, human capital, urbanization and initial education inequality seem to be more important in explaining access to education. Any

statistical significant differences across countries depending on their development- and democracy level are not found.

This paper starts with a brief presentation of relevant concepts. Thereafter it continues with a description of the hypotheses the study is built upon. The next following chapter describes the empirical specification together with a presentation of the data. In the next section the empirical analysis is to be found, it begins with a brief descriptive analysis followed by results from the baseline regression. This section also covers the sensitivity analysis where the robustness of the results is tested. This study finishes of with some concluding remarks.

## 2. BACKGROUND

### 2.1 Access to education

Access to education is affected by both supply and demand factors. Regarding the supply side, the financial means available to the educational system are of great importance. These expenditures vary considerably among countries, the per-student public expenditure in PPP dollars (Purchase Price Parity) is reported in *Table 1*.

*Table 1. Shows the per-student public expenditure over low, middle and high income countries.*

Type of country	Primary education	Secondary education
Low-income	202	366
Middle-income	833	1013
High-income	3059	3915

(Holsinger and Jacob, 2008, p. 91)

The decision of resource allocation between and within sectors is often a political decision. These decisions are optimally taken with information about the costs, benefits and effectiveness of interventions together with concern for equity. However, these decisions are often influenced by political considerations and social attitudes. Wu exemplifies this by stating that considerations are often taken to the “most vocal constituencies” who often receive the “lion’s share of resources and attention” (Wu, 2008, p.93), whereas rural areas, poor regions or regions where ethnic and religious minority lives, seldom are prioritised as these groups often are less powerful in a society.

Access to education is, as an effect of limited financial resources, also constrained by insufficient supply of schools, unqualified teachers and scarcity of learning resources. With long distances to school and no transportation, children will not be able to attend school until they are able to walk the distance by themselves and if the way is judged too risky this affects the child's participation in school. In case of undersupply of teachers the pupil-to-teacher rate becomes high, pressuring the education to be held in shift with few hours of education per pupil as a result. High pupil-to-teacher rate also lowers the possibility for the teachers to give the necessary help and support to all the pupils. According to Wu, the most severe problem is institutional constraints, referring to teacher management and teacher accountability as well as incentives. In many developing countries with low levels of secondary and tertiary education there is shortage of qualified teachers. This is more severe in the rural areas as it is difficult to attract teachers to take positions there, and in some minority communities there are no qualified teachers who can speak the appropriate language. There are also studies showing troublesome high absence of teachers. (Wu, 2008, pp. 93-94)

There are several demand side constraints affecting access to education as well. For poor people the cost of sending their child/children to school comes with a trade-off between a future greater earnings possibility for the child and an additional immediate cost to a maybe already stressed budget. Even if the education is tuition free it will be accompanied with cost for uniforms, books etc. Poor families will often also face the cost of education in terms of opportunity cost of the lost labour supply from the child. Other factors affecting the demand for education are social exclusion and discrimination which to a larger extent affect girls and minority groups. (Wu, 2008, pp. 94-96)

### ***2.1 Inequality of opportunity***

Four thinkers, John Rawls, Amartya Sen, Ronald Dworkin and John Roemer have in particular contributed to the equality debate. They have contributed in different ways, however, all four thinkers' work have lead the social justice debate to a larger extent to focus on the distribution of opportunities and not the distribution on particular outcomes. (World Development Report 2006, p. 77) To present the thoughts about the concept of inequality of opportunity, this study focuses on the work of Roemer.

Roemer argues that there are two concepts for equality to be present. The first is what can be called "level-the-playing-field". It refers to the intervention of the society to give people with different backgrounds more equal starts-offs and is most relevant in young ages, during the period of formation. The second concept is called "non-discrimination principle".

It refers to the practice of meritocracy; that individuals are judged only by relevant attribute. The first of these concepts is the one most relevant to the discussion of equality of opportunity. (Roemer, 1998, pp. 1-12)

An individual's achievements in different aspects of life, such as education but also socio-economic condition, will be determined by both internal and external resources. Resources counted as internal are for example; genes, neighbourhood, parents' education and effort etc., while examples of external resources can be; education, training, teachers and books etc. The thought behind "level-the-playing-field" is to compensate for unequal ability to produce/achieve education or social-economic condition by the means of external resources. However, this principle of "level-the-playing-field" should only be accountable to resources determined outside the individuals' own control, called circumstances. However, differences in outcome not due to circumstances, but due to effort taken which is controlled for by the individual itself, should not be compensated for. The question then becomes what should be accounted for as circumstances in an individual's life and how to be able to measure these different variables. (Roemer, 1998, pp. 1-12)

Roemer gives a suggestion to a policy of equality of opportunity. He suggests that individuals should be divided into different types depending on their specific circumstances (decided in a society what to be accounted as circumstances). Within these types, individuals will be ranked according to the degree of effort they take on and not the absolute level, since this will be determined by the individual's circumstances to some extent. The equality of opportunity policy is then to find the distribution of resources among the types so that the individuals in the same e.g. centile in different types will be able to reach the same achievements/outcome. (Roemer, 1998, pp. 1-12)

### ***2.3 Inequality in education***

During the period 1960-2000, global inequality in education developed in the desirable direction. As *Figure 2* shows, this development was accompanied with almost a doubling of the mean years of schooling.



There are several different dimensions or groups to compare within countries, for example; rural – urban, wealthy – poor, educated – non-educated parents and regional differences. What is of great concern regarding inequality is that it tends to perpetuate and reinforce itself over time and generations. The authors of World Development Report from 2006 emphasize that different dimensions of inequality such as in health, education, income and power are often correlated. An individual's level of education will affect its future income, its health and the health of its children as well as its ability to interact and communicate with others and its political power. Just as an individual's economic well-being will affect its educational- and health outcome. (World Development Report 2006, pp. 28-43)

#### ***2.4 Inequality in income***

Just as with education inequality, income inequality can be decomposed into between- and within-country inequality. However, a difference between the two inequalities is found regarding the share among them. As previously mentioned, less than one-fifths of the inequality in education is due to between-country inequality. Regarding income inequality, the opposite is valid. Firebaugh reports that between two-thirds and three-fourths of the total income inequality is due to between-country inequality. This is not due to within-country inequality being small but to the fact that between-country inequality is large. This, as the average income of the richest country is 30 times the average income in the poorest country. (Firebaugh, 2000, pp. 329-336) Just as with inequality in education, this study uses within-country inequality for income as well. Human Development Report from 2010 states that within-country income inequality has since the 1980s increased. Countries report differences in their development in income inequality, but for every country succeeding in reducing income inequality, two countries experience a worsening. Countries in the former Soviet Union have during the transition period been especially represented among the countries whose inequality has worsened. This is also valid for former centrally planned countries like China and Vietnam. This trend is also seen along most countries in East Asia and Pacific, where increases in income inequality are mainly due to growing differences between rural and urban areas. Latin America and the Caribbean experienced high income inequality due to prevailed inequality in land ownership and education level. However, several countries in this region, have through progressive policies succeeded in turning this development in the desired direction. Sub-Saharan Africa has also to some extent experienced a reduction in income inequality since the 1990's.

### 3. HYPOTHESES

There are three channels through which income inequality can affect the outcome of education, namely; expenditure, enrolment and quality. Regarding expenditure and resources to the educational system, there are theories suggesting that inequality affects the level of redistribution in a society where part of the redistribution can be to support public education. Theory also suggests that inequality of income will affect the enrolment rate of children in education through the subsistence constraints faced by families. The quality of education and thereof also the utility benefitted from taking part of education are given by the expenditure per student. As inequality affects both expenditure and enrolment rates in education and hence the resources per student, inequality of income also affects the quality of education. (Gutierrez-Sourdis, 2006, pp. 1-2, Gutierrez and Tanaka, 2009)

This calls for a further examination of how income inequality affects access to education. This study will explore four different hypotheses regarding the relationship between inequality in education and inequality in income. The first two hypotheses consider the possibility and choice among families to send their children to school. While the first hypothesis considers a subsistence constraint, the second hypothesis takes into account the calculations of net benefits from education. The last two hypotheses concern the issue of redistribution and how the level of redistribution will be affected by the income distribution. The level of redistribution, can in turn, be thought of to be affecting the access to education and hence inequality in education. As evident, the hypotheses have different channels through which the income distribution affects access to education, and they are also inconclusive in what direction the relation goes.

The first hypothesis takes into consideration the *possible subsistence constraint household might be facing*. At a given GDP level, high income inequality results in a larger number of poor people who tend to send their children to school to a lesser extent. Theoretically, this is explained by the subsistence constraint. Households face an immediate cost of sending their child to school even if education is freely provided. This is due to the associated opportunity cost of either child-labour income or household shores. For households this can be a non-trivial income that determines the welfare of the household. It is assumed that below a certain income, families are unable to send their children to school whereas families above that income do send their children to school. This hypothesis predicts a negative relationship between income inequality and access to education. (Gutierrez-Sourdis, 2006, p. 3, Gutierrez and Tanaka, 2009)

However, as Gutierrez-Sourdis mentions, there is evidence that, even among families above the subsistence constraints, child-labour is non-trivial. This evidence gives rise to the second hypothesis regarding the relation between income inequality and education inequality. When the subsistence constraint is being relaxed, the parents' decisions to send their child to school or not is instead *based upon the net benefit from acquiring education*. The net benefit depends on the cost, including the opportunity cost, and benefits from higher future earnings. This means that the threshold determining if parents decide to send their child to school or not will be dependent on the level of income inequality in the society. As in a highly unequal society the benefits of acquiring an education will be greater. This hypothesis predicts a positive relationship between income inequality and access to education. (Gutierrez-Sourdis, 2006, p. 3, Gutierrez and Tanaka, 2009)

The median voter hypothesis, developed by Hotelling (1929) and Downs (1957) and later extended by Roberts (1977), suggests that the median voter will be the decisive vote in an election. This result is dependent upon the assumptions of majority rule, high participation of voters, and that preferences are single peaked. This hypothesis is later applied to the relation between inequality and redistribution in a society (Meltzer and Richard 1981). Meltzer and Richard claim that the size of the government, that is the share of income redistributed by government, depends on the relation between the mean income and the income of the decisive voter. In societies with unequal income distribution, the distribution will be skewed to the right which implies that the mean income lies above the median income. In more unequal societies the individual with the median income, that is the median voter, will be relatively poorer to the mean income. This implies that in a progressive taxation system *the more unequal the income distribution, the more the median voter can gain from redistribution*. This results in a situation where societies with more unequal income distribution will tend to vote for higher redistribution.

This gives rise to a third hypothesis regarding the relationship between income inequality and access to education. Greater redistribution can be assumed to result in more resources to the educational system which will lead to a situation where more supply side constraint to access to education will be relaxed. Moreover, if these increased resources are assumed to raise the quality of the education and hence also the benefit from attaining education, this will lead to more parents sending their children to school. This hypothesis predicts a positive relationship between income inequality and access to education.

A fourth hypothesis over the relationship between income inequality and access to education is found from another model considering the redistribution of resources. The

model, developed by Alesina et al. (1999) suggests that *heterogeneity of preferences* among ethnic groups affects the level of public goods. Their empirical analysis of US cities shows that “voters choose lower public goods when a significant fraction of the tax revenues collected on one ethnic group are used to provide public goods shared with other ethnic groups.” (Alesina et al., 1999, p.1244) The model is built upon the median voter theorem in accordance with the assumption that voters first vote over the tax level and thereafter on the type of public good. It is also assumed that the voters are fully aware that after the level of taxation is chosen the type of public good will be the one preferred by the median voter. In a highly polarized society this yields that the median distance from the median voter is larger. It results in a situation where a large fraction of the population has preferences far away from the median voter and therefore prefers a low taxation as the median vote is the decisive voter over the type of public good to use the resources on. This means that in equilibrium, the amount of public goods is decreasing with polarization. (Alesina et al., 1999, pp 1243-1251)

This model can be applicable to other types of heterogenic preferences in a society. Alesina et al. mention that polarization of preferences can also be due to polarization of income. (Alesina et al., 1999, p. 1257) In accordance with the previous hypothesis, a low level of redistribution is assumed to result in fewer resources to education. Fewer resources to the educational system can be assumed to aggravate supply side constraints in access to education. It is also assumed that fewer resources to education will deprive quality and hence result in a situation where it is not as beneficial to acquire education. This hypothesis predicts a negative relationship between income inequality and access to education.

*Table 2. Summary of hypotheses and the belonging predicted sign of the relation between income inequality and access to education and income inequality and education inequality.*

<b>Hypothesis</b>	<b>Predicted sign on access to education</b>	<b>Predicted sign on inequality in education</b>
Subsistence constraint	Negative	Positive
Net benefit from education	Positive	Negative
Inequality and redistribution	Positive	Negative
Heterogeneity of preferences	Negative	Positive

### ***3.1 Development and democracy level***

The first two hypotheses predict a relationship between income inequality and education inequality that works via the cost of education. How well these two hypotheses are applicable to countries might be a question of the countries’ development level. In low income countries,

the risk of a larger part of the population being poor is evident, especially if the income distribution is uneven. If poverty is more extensive, it is more reasonable to believe that the opportunity cost of sending children to school is higher and also subsistence constraint is being binding. This can, to lesser extent be thought of as being an issue in high income countries. Therefore, it can be presumed that the hypothesis regarding the subsistence constraint and the hypothesis regarding net benefit from acquiring an education is more applicable to low income countries.

For the last two hypotheses the relationship works via redistribution, as is rather distinct from the two hypotheses just mentioned. First, for these two hypotheses to be applicable, a country needs to have a working voting system, which can be assumed to be more often present in democracies. These hypotheses also assume a functioning redistribution through a taxation system. This feature can more often be found in more developed countries. Therefore, the hypothesis regarding inequality and redistribution and the hypothesis regarding heterogeneous preferences can be assumed to be more applicable to democracies and high income countries.

#### **4. PREVIOUS RESEARCH**

Few empirical studies concerning determinants of access to education can be found. Three studies with to some extent various results will be presented here. The first is a study by Susan Mayer, who investigated the effect of changes in income inequality between 1970 and 1990 on the mean educational attainment between States in the US. She found that the overall years of education increase and this is due to increased college enrolment. This result is only partly due to the incentives of higher returns to education. Beside this, she also found that the increase in income inequality lead to an increase in the inequality in educational attainment between rich and poor. While an increase in income inequality increased the educational attainment among rich children, it actually decreased the educational attainment among poor children. (Mayer, 2001, pp. 14-17)

Gutierrez-Sourdis studies the impact of income inequality on enrolment rates in education for 59 developing countries over the period 1960-1990. Level of development and various demographic variables are added as control variables. The empirical study shows a negative impact of income inequality on gross enrolment rates in both primary and secondary education. GDP per capita is shown to be positive and concave. This study controls for regions and the only region turning out positive is Sub-Saharan Africa. The variable, fraction of population under 15 years, as a proxy for number of children per household, is positive and

significant for secondary education only, whereas radio per 1000 people, as a proxy for density and urbanization, is found to be positive for both primary and secondary education. (Gutierrez-Sourdis, 2006, pp. 21-23)

A third study, expanding both the number and types of countries, is conducted by Checchi. His study looks at the impact of income inequality on gross enrolment rates in education over the period 1960-1990 for 102 countries. In this study, imperfect capital markets are assumed and hence education is wealth constrained and a negative relation between income inequality and enrolment in education is expected to be found. Checchi controls for variables affecting enrolment to education both from households' demand for education and from government supply of education. The level of development is also controlled by the GNP per capita. Checchi does not find evidence that primary and higher education is constrained by wealth. The study however finds a negative statistically significant relation for enrolment in secondary education. However, when controlling for gender it is found that female children are financially constrained even at primary education level and also at a larger magnitude. Development level seems to be related in a positive way to enrolment rate for primary, secondary and higher education. Labour demand for skilled workers, proxied by ratio of physical capital stock to GDP, seems to play a role for enrolment ratios in secondary and higher education. The study finds some evidence however, not strong, of educational expenditure affecting the enrolments rates. From the demand side, the variable over child mortality as a proxy for poverty, has a negative significant impact on primary education. Fertility rate, as a proxy for number of children in a family, shows a positive effect for primary education which illustrates evidence of a support effect within families. While for secondary education, this variable is shown to have a negative impact on enrolments rate, which can be evidence of a resource constraint effect. (Checchi, 1999, pp. 16-23)

To sum up, there seems to be indication of a negative relation between income inequality and education enrolment. Mayer however finds a positive relation, but as she controls for household income, the positive relation is only valid for rich children. For poor children the relationship between income inequality and educational attainment is negative. Development levels have shown to be positively related to enrolment rates in both studies including that variable. There is indication of a positive relation between urbanization and enrolments rates whereas there is indication of a negative relation between poverty and enrolment. Number of children in family has shown relations going in different directions. As these studies show indications of a direct affect of income inequality on access to education

there are likely several indirect ways for income inequality to effect access to education. There is indication that income inequality affects individuals' health in a negative way. (Wilkinson and Pickett, 2005) There is also indication of income inequality affecting growth, however, the sign of the relation is still an issue for debate (Barro, 2000). Studies has also shown that unequal income distribution increases social and political instability (Alesina and Perotti, 1996) In turn, it is reasonable to believe that both individuals' health, an economy's growth and social and political instability will affect access to education.

## 5. EMPIRICAL SPECIFICATION

To investigate the relation between education inequality and income inequality the following model is formulated:

$$\Delta edu. ineq_i = \alpha' X_i + \beta \Delta inc. ineq_i$$

where  $\Delta edu. ineq$  represents the change in education inequality between 1980 and 2000,  $X$  is a matrix of control variables and  $\Delta inc. ineq$  represents the change in income inequality between 1975-1995,  $i$  is the indexation of countries.

This study follows Sylwester (2002) in his approach, by regressing the change in the dependent variable on an explanatory variable and a set of control variables. However, this study extends that method by also using the change in the explanatory variable. The advantage with applying this method is that it deals with the issue of endogeneity; particularly time invariant heterogeneity and omitted variables. Omitted variables can otherwise cause the coefficient to be biased as it captures the indirect effect from omitted variables that are affected by income inequality that in turn affect education inequality. This problem is also mitigated by including various control variables in to the model that can be assumed to affect access to education. By using the change in both the dependent and the independent variables, the model also controls for time invariant variables to some extent. Time invariant variables control for differences between countries that can affect the relation investigated, that are constant over time.

By using the change in both education inequality and income inequality this study hopefully will be able to say something about the causal effect of income inequality on access to education to a larger extent.

### ***5.1 Dependent and independent variables***

The aim of this study is to empirically investigate how income inequality affects access to education. While access to education is not a straight forward concept to measure, a couple of

possible indicators can be thought of. Gutierrez-Sourdis and Checchi chose to use enrolments rates in their studies while Mayer chooses to use mean educational attainment and overall educational attainment. This study however, uses the Gini coefficient of education as a proxy for access to education. The choice is based upon the fact that this study also aims to incorporate the concept of inequality of opportunity. For this reason, the Gini coefficient is a more reasonable measure to use in this study as it incorporates the distribution of education among the population.

The Gini coefficient is a common measure of inequality. It is proportional to the Lorenz-curve and gives a measure of the area between the Lorenz-curve and the 45-degree line symbolizing perfect equality. It measures the distribution of a particular resource (for example; income, wealth and land) among a population as each point on the Lorenz-curve measures the share of the resource given to a particular share of the population. The Gini coefficient ranges between zero and one. Zero indicates perfect equality as the Lorenz-curve and the 45-degree line perfectly coincide. When the distribution is not perfectly equal the Lorenz-curve starts to fall below the 45-degree line, the area between this line and the Lorenz-curve starts to increase and hence the Gini coefficient increases. A Gini coefficient of one, indicates perfect inequality where the Lorenz-curve runs parallel with the y-axis until 100% of the population, where it intersects with the Lorenz-curve.<sup>1</sup>

The explanatory variable, income inequality, is also measured by the Gini coefficient. This study also follows Sylwester (2002) by using lagged values for the explanatory variable, using the change during 1975-1995 for income inequality while using the change during 1980-2000 for education inequality. This is due to the belief that it takes some time for income inequality to affect access to education. By lagging the explanatory variable the problem of reverse causality is to some extent mitigated. This is an important issue to consider as it is possible to think that not only income inequality affects education inequality, but it is also reasonable to imagine that education inequality affects income inequality.<sup>2</sup>

The change in inequality in education and income are both measured such that a positive number of this index denotes an increase in inequality.

---

<sup>1</sup> For a more technical description of the Gini coefficient see Thomas et al. (2001).

<sup>2</sup> Dao (2008) empirically test the source of income inequality and finds education inequality to be influential.

## ***5.2 Control variables***

As mentioned earlier, several control variables are included in the model, represented by the matrix; X. This matrix is comprised of: the natural logarithm of real GDP 1975, change in average years of education 1995-1975, percentage of total population living in urban areas 1975 and education inequality 1975. All these variables are included in the model as they are believed to affect access to education in different ways.

First, GDP per capita is included in the model, to control for the level of development. It is reasonable to believe that an initial higher level of GDP per capita would mean that the country is better equipped to create a more equal distribution in educational attainment. This would predict a negative relationship between the change in education inequality and initial GDP per capita. However, it can also be reasonable to assume that in countries with an initial higher GDP per capita the educational distribution was rather equal in 1980 as most individuals obtained the same level of education. In 2000 however, as a larger share of the population got the opportunity to obtain a higher education, the distribution of educational attainment has become more unequal. The opposite could be believed to have occurred in countries with low initial GDP per capita. In 1980 the education distribution was highly unequal with a large share of the population not obtaining any education and a smaller part obtaining some education whereas in 2000, a larger share of the population had obtained some basic level of education. This development results in the distribution of educational attainment becoming more equal. This reasoning would predict a positive relationship between the change in education inequality and initial GDP per capita. Hence the predicted sign of the coefficient for GDP per capita is inconclusive.

This study also includes the change in average years of education in the model. It is quite reasonable to believe that if a country manages to increase the average years of education that development will have an effect on the distribution in educational attainment. Previous research has shown a negative relationship between the average years of education and the Gini coefficient in education. (Thomas et al., 2001, p. 18) This gives reason to predict a negative sign on the coefficient of change in average years of education.

Another control variable included in the model is the share of urban population in a country, which can be thought of to affect inequality of education in a number of ways. It is reasonable to believe that access to education is easier to achieve in urban areas where population density is high. This would yield a negative relationship between urbanization and the change in education inequality. However, it can also be reasonable to believe that rural areas are more equal but also poorer. The first condition would predict a positive relationship

between the share of the population that live in rural areas and the change in education inequality whereas the second condition predicts a negative relationship among those two variables.

Finally, initial level of education inequality 1975 is included in the model. The reason to include this variable is to control for the possibility that countries with initial low inequality in education are “bounded from below”, this as it is difficult or impossible to make further large improvements in the education distribution whereas countries with initial higher inequality have a larger room for a change to take place.

### 5.3 Data

To be able to empirically investigate how income inequality affects access to education, the above cross-country model is used. When organising an appropriate dataset to use, a few issues have to be considered. First, the dataset used for this study contains several variables which are collected from different sources. However, it is important that each variable has compatible data. In this study, it is a relevant issue to deal with, as both education- and income inequality data need two observations for each country as the change in the variable is used in the regression. This is an important issue to consider as, Sylwester (2002) mentions, different inequality dataset can be obtained in dissimilar ways. In case of using incompatible dataset, it cannot be known to what extent the change in inequality is due to an actual change or just due to dissimilarities in measurement techniques. In order for this not to become a problem, in this study the completed dataset, with 64 country observations, was dependent upon the possible collection of inequality data for both education and income.

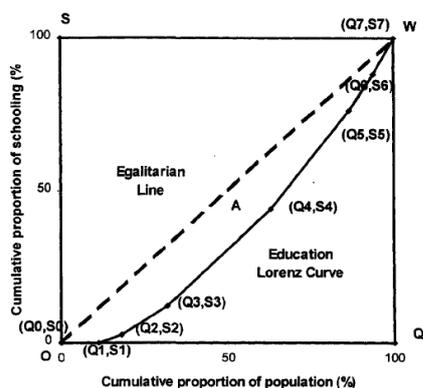
The data for education inequality is taken from a dataset constructed by Thomas, Wang and Fan. It was developed in 2001 and covers observations for 105 countries over the period 1960-1990. The Gini coefficient is based on data over educational attainment of the population above 15 years. To get some intuition of the method behind the Gini coefficient of education *Figure 4* shows an education Lorenz curve. Thomas et al. use seven levels/categories for different education attainment, which are listed in *Figure 5*. Moreover, explanations for how the Lorenz curve in *Figure 4* is calculated can also be found in *Figure 5*.

3

---

<sup>3</sup> For a more technical description of the calculations of the Education Gini coefficient see Thomas et al. (2001).

Figure 4. Illustration of the Education Lorenz Curve.



(Thomas et al., 2001 p. 12)

Table 3. Table over characteristics in calculation of Education Lorenz Curve.

Levels in attainment	Prop. of pop.	Years of edu.	Cum. prop. of pop.	Cum. Prop. of edu.
Illiterate:	$p_1$	$y_1 = 0$	$Q_1 = p_1$	$S_1 = (p_1 y_1) / \mu = 0$
Partial-Primary:	$p_2$	$y_2 = y_1 + 0.5C_p$	$Q_2 = p_1 + p_2$	$S_2 = (p_1 y_1 + p_2 y_2) / \mu$
Complete-Primary:	$p_3$	$y_3 = y_1 + C_p$	$Q_3 = p_1 + p_2 + p_3$	$S_3 = (p_1 y_1 + p_2 y_2 + p_3 y_3) / \mu$
Partial-Secondary:	$p_4$	$y_4 = y_3 + 0.5C_s$	$Q_4 = p_1 + p_2 + p_3 + p_4$	$S_4 = (p_1 y_1 + p_2 y_2 + p_3 y_3 + p_4 y_4) / \mu$
Complete-Secondary:	$p_5$	$y_5 = y_3 + C_s$	$Q_5 = p_1 + p_2 + p_3 + p_4 + p_5$	$S_5 = (p_1 y_1 + p_2 y_2 + p_3 y_3 + p_4 y_4 + p_5 y_5) / \mu$
Partial-Tertiary:	$p_6$	$y_6 = y_5 + 0.5C_t$	$Q_6 = p_1 + p_2 + p_3 + p_4 + p_5 + p_6$	$S_6 = (p_1 y_1 + p_2 y_2 + p_3 y_3 + p_4 y_4 + p_5 y_5 + p_6 y_6) / \mu$
Complete-Tertiary:	$p_7$	$y_7 = y_5 + C_t$	$Q_7 = p_1 + p_2 + p_3 + p_4 + p_5 + p_6 + p_7$	$S_7 = (p_1 y_1 + p_2 y_2 + p_3 y_3 + p_4 y_4 + p_5 y_5 + p_6 y_6 + p_7 y_7) / \mu$

The income distribution data used in this study is taken from the Standardized World Income Inequality Database. It covers observations from 1960 to 2008 over a total of 150 countries. When constructing the dataset used in this study, consideration had to be taken concerning which countries to include and over what period the change should be taken. The aim was to get a dataset with as many observations as possible.

Another important consideration when using changes in a variable is to derive the change over a sufficient long period for a change in the distribution to occur. Also in this regard, this study follows Sylwester (2002) and derives the change over a 20 years period in both inequality variables.

The dataset over education inequality is almost complete regarding observations over all points in time for all countries. However, the data set over income inequality is complete to a lesser extent and especially for earlier observations. Hence, the dataset over income inequality determined over what countries to include in the dataset and over what years to derive the change. In order to extend the dataset as much as possible, the observations over income inequality which are taken over the period 1995-1975 are in case of missing value for those particular years replaced by a nearby year, the limit is set at five years away.

## 6. EMPIRICAL ANALYSIS

### 6.1 Descriptive analysis

For the countries included in this study a few interesting features are worth highlighting. Regarding the levels in education inequality, it can be observed that the mean is somewhat smaller in 2000 than in 1980, reduced from 0.37 to 0.30. For income inequality the mean is stable for the two time points, 1995 and 1975 at 0.39. Norway, US and Poland are the countries with the lowest education inequality in 2000, the three countries with the highest inequality are: Senegal, Nepal and Sierra Leone. The countries with lowest income inequality in 1995 are: Finland, Denmark and Sweden whereas the countries with the highest income inequality are: South Africa, Malawi and Zambia. It is also worth noting that the range is larger in education inequality compared to income inequality. Norway has an education Gini coefficient of 0.087 and the corresponding number for Sierra Leone is 0.724, whereas Finland has an income Gini coefficient of 0.215 and the corresponding number for Zambia is 0.562.

As become evident by *Figure 5* more than two thirds (46/64) of the countries have succeeded in reducing the Gini coefficient in education. Among the countries succeeded the best in reducing the Gini coefficient in education most countries are from Sub-Saharan Africa, Middle East and North Africa and South Asia. Among the countries that showed a positive change, i.e. a worsening of the education inequality most countries are from Europe and Central Asia, Latin America and East Asia and Pacific. Arab Republic of Egypt presents the largest improvement in the inequality in education with a drop in the education Gini coefficient from 0.753 to 0.457, a reduction of 0.296 units. Islamic Republic of Iran and Uganda also carried out large reductions in their Gini coefficients with a reduction of 0.224 units. Regarding the change in income distribution fewer countries, less than half the sample (29/64), succeeded in reducing the Gini coefficient.

All countries have succeeded in increase the average years of education between 1975 and 1995. The mean increase in average years of education is almost one year (0.92). The smallest increase was carried out by Tanzania with a modest increase of 0.06 years. Norway carried out the largest increase with an increase of the average years of education of over four years.

Table 4. Descriptive statistic.

Variable	Explanation	Mean	Std. Dev.	Min	Max	n	Source
Δ Education Gini	Change in education inequality 2000-1980 (Gini coefficient of educational attainment)	-0.065	0.08	-0.296	0.038	64	Thomas et al., 2001
Δ Income Gini	Change in income inequality 1995-1975 (Gini coefficient of net incomes)	0.004	0.048	-0.138	0.102	64	Solt, 2009
GDP per capita	Natural logarithm of real GDP per capita (PPP adjusted)	8.618	0.963	6.52	9.98	64	Heston et al., 2009
Education Gini 1975	Education inequality 1975 (Gini)	0.402	0.211	0.09	0.953	64	Thomas et al., 2001
Δ Human capital	Change in average years of education 1995-1975	1.766	0.92	0.06	4.16	64	Thomas et al., 2001
Urban population	Urban population 1975 (percentage of total)	48.614	24.87	4.8	100	64	World Bank, 2010a
Democracy level	Democracy level (Polity IV variable 1975)	1.25	7.974	-10	10	64	Marshall et al., 2009
Democracy dummy	Democracy (+6 - +10 in Polity IV variable 1975)	0.469	0.503	0	1	30	Marshall et al., 2009
High income dummy	High or upper middle income countries 1975 (up to \$3,945 GNI/capita, GDP used here)	0.672	0.059	0	1	43	Constructed
Low income dummy	Low or lower middle income countries 1975 (from \$3,946 GNI /capita, GDP used here)	0.328	0.059	0	1	21	Constructed
Population < age 14	Population under age 14 1975 (percentage of total)	6.203	3.879	3	15	64	World Bank, 2011
Population > age 65	Population over age 65 1975 (percentage of total)	36.859	9.083	20	50	64	World Bank 2011
LA and C	Latin America and the Caribbean	0.219	0.417	0	1	14	World Bank, 2010b
SA	South Asia	0.078	0.27	0	1	5	World Bank, 2010b
SSA	Sub-Saharan Africa	0.172	0.38	0	1	11	World Bank, 2010b
EA and P	East Asia and Pacific	0.172	0.38	0	1	12	World Bank, 2010b
ME and NA	Middle East and North Africa	0.078	0.27	0	1	5	World Bank, 2010b
E and CA	Europe and Central Asia	0.234	0.427	0	1	15	World Bank, 2010b
NA	North America	0.031	0.175	0	1	2	World Bank, 2010b

Figure 5. Chart displaying the change in education inequality 2000-1980 for all countries in the data set.

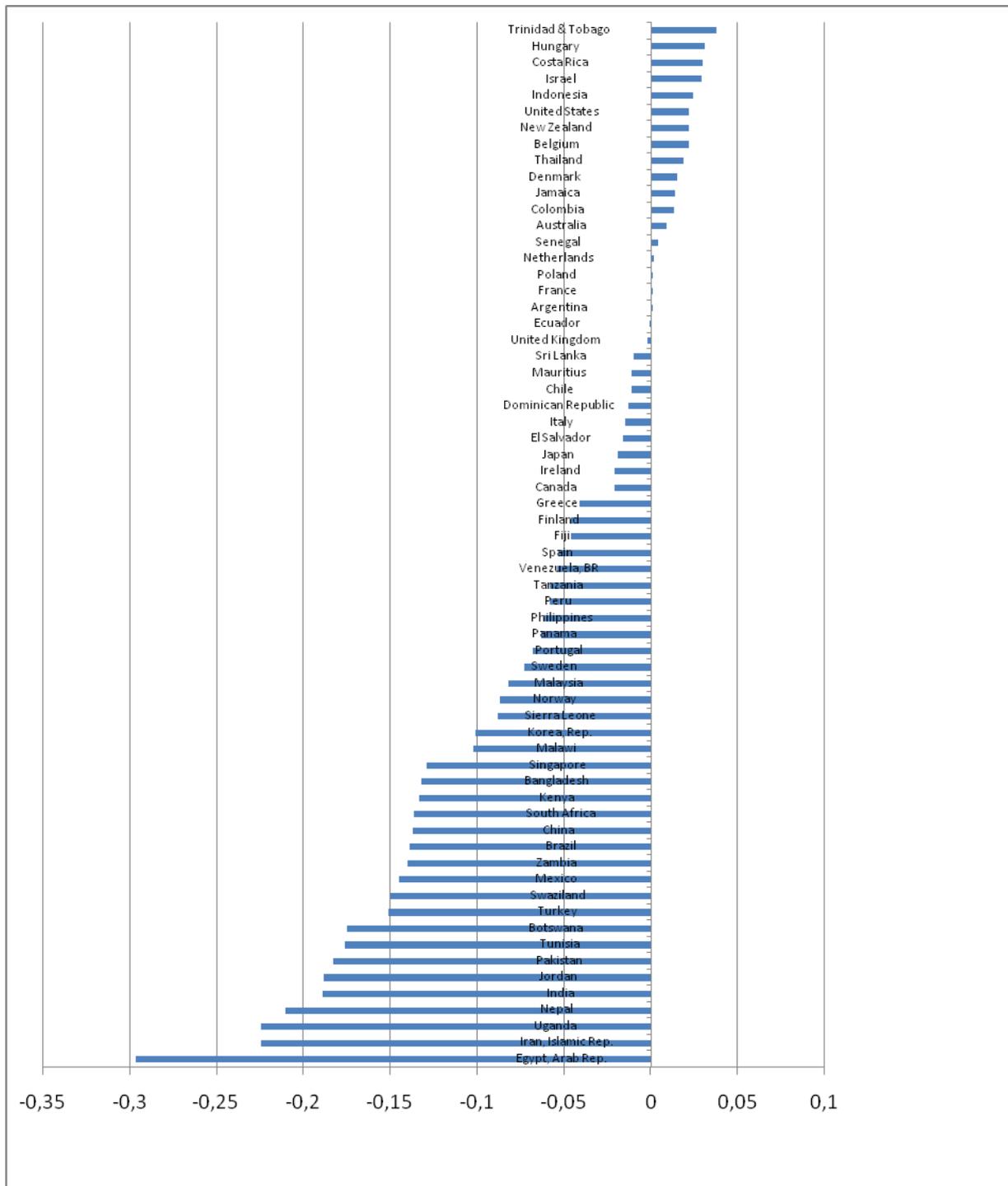


Figure 6. The cross-country correlation between education inequality 1975 and income inequality 1980 divided between low or lower middle income countries (left) and high or upper middle income countries (right).

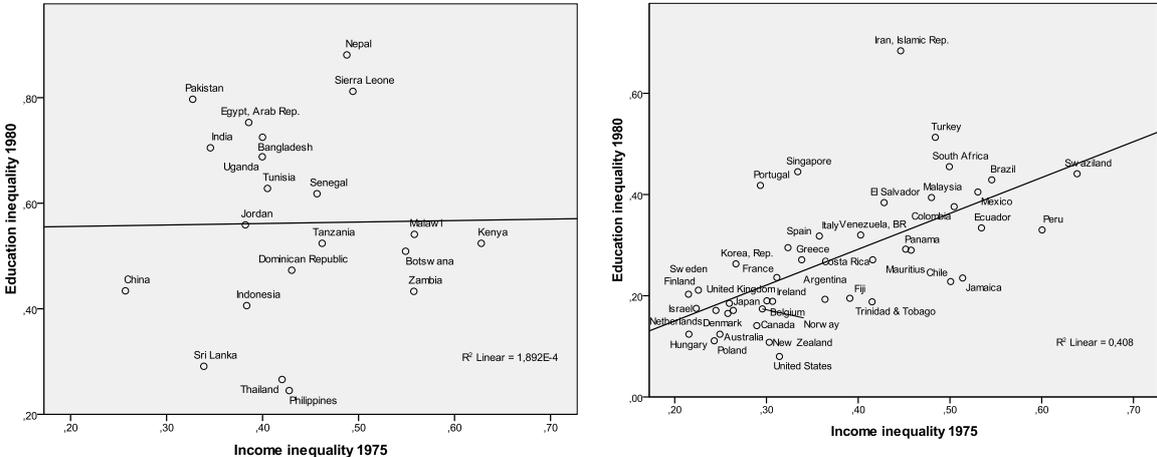


Figure 7. The cross-country correlation between education inequality 1975 and income inequality 1980 divided between non-democracy countries (left) and democracy countries (right).

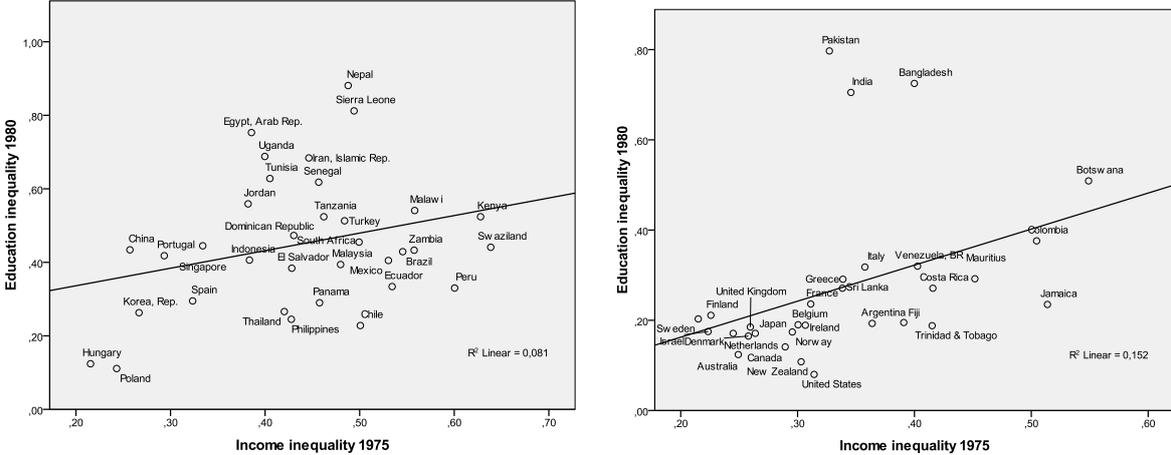


Figure 6 and Figure 7 show correlations between education inequality and income inequality divided across low and high income countries as well as across democracy and non-democracy countries. As evident from Figure 6 the slopes for low or lower middle income countries and high or higher middle income countries are rather different from each other. The correlation for low or lower middle income countries are almost showing a negative relation. For democracy and non-democracy countries both slopes are positive but a clear distinction in magnitude of the slopes can be seen. This gives reason to investigate these differences further which will be done in the sensitivity analysis.

## 6.2 Regression analysis

The estimated results from the OLS- regression are presented below in *Table 5*. The model is built up according to the specific-to-general method.

As can be seen from the simplest model-specification presented in column one, there is a positive relation between income inequality and education inequality that is significant at a 5% level. In this model, a one unit reduction in income inequality would yield a 0.5 unit reduction in education inequality. Kenya is the country in the sample used in this study with the largest reduction of income inequality between 1975 and 1995, with a reduction of 0.138 units. A reduction in income inequality as presented by Kenya would hence yield a reduction of 0.07 units in education inequality.

*Table 5. Regression results from baseline regression.*

$\Delta$ Education Gini	(1)	(2)	(3)	(4)	(5)	(6)
$\Delta$ Income Gini	0.500** [0.203]	0.340* [0.184]	0.260* [0.155]	0.291* [0.155]	0.060 [0.125]	0.092 [0.124]
GDP per capita 1975		0.040*** [0.010]	0.041*** [0.008]	0.056*** [0.013]	0.004 [0.008]	0.020* [0.012]
$\Delta$ Human capital			-0.041*** [0.008]	-0.042*** [0.008]	-0.032*** [0.006]	-0.033*** [0.006]
Urban population				-0.001 [0.001]		-0.001* [0.004]
Education Gini 1975					-0.25*** [0.040]	-0.252*** [0.040]
Constant	-0.067*** [0.010]	-0.403*** [0.080]	-0.350*** [0.070]	-0.443*** [0.100]	0.060 [0.083]	-0.044 [0.100]
Number of countries	64	64	64	64	64	64
Adjusted R-squared	0.074	0.275	0.492	0.499	0.689	0.702

\*, \*\* and \*\*\* denotes statistical significance at 10%, 5% and 1% levels, respectively.

Standard errors in brackets.

All model-specifications in *Table 5* yield a positive relationship between income inequality and education inequality which is in accordance with two hypotheses. The hypothesis based on *subsistence constraint* and the hypothesis based *heterogeneity of preference* both predict a positive relation between income inequality and education inequality. This gives some

indication that these two hypotheses are more relevant in explaining the relationship between income inequality and education inequality than the remaining two hypotheses.

However, as can be seen from *Table 5*, as the model is extended to include control variables, both the significance and the magnitude of the coefficient of the change in income inequality diminish. The relationship between the two variables in focus in this study becomes insignificant and also decreases drastically in magnitudes as initial level of education inequality is introduced to the model. A high correlation between the change in income inequality and education inequality 1975 could be suspected, but as can be seen from the correlation matrix in *Appendix 2*, the correlation between the two variables correlates at reasonable -0.33.

In the most developed model, presented in column six, all control variables included are statistical significant. Hence, it appears as access to education is not determined by income inequality but rather by initial GDP, human capital development, urbanization and initial level of education inequality.

The coefficient on initial GDP per capita variable shows a positive relation, significant at 10% level. An increase in initial GDP per capita with 10% yields a 0.002 increase in the education Gini coefficient during the period 1980-2000. This means that countries with an initial higher GDP per capita worsened their inequality in education over the period.

The largest coefficient is found for the initial level of education inequality, at -0.252, which is significant at 1% level. This shows that countries with higher inequality in education 1975 have to a larger extent reduced inequality in education during 1980-2000, which was an expected result.

Regarding urbanization, for which theory does not present consensus on the sign of the relationship, the coefficient turns out to be negative. This means that countries with higher level of urbanization experienced a larger reduction in inequality in education. The change in average years of education 1975-1995 also has a negative relation to the change in education Gini. This means that a country succeeding in increasing average years of education also has been able to reduce the education inequality. This is in line with what previous research has shown.

### 6.3 Sensitivity analysis

Several additional control variables were included in the analysis to control the robustness of the result from the baseline regression results. These results are reported in *Table 6* and *Table 7*. The data over the control variables are also taken from 1975, in order to decrease the risk of reverse causality.

As the result from the baseline regression showed, the relation between income inequality and education inequality was always positive. However, as discussed in the chapter presenting the four hypotheses, it can be suspected that the above mentioned relationship goes in different directions depending on the countries level of development as well as level of democracy. This issue became further stressed by the correlations presented in *Figure 6* and *Figure 7*. It therefore seems important to investigate the possibilities of different magnitudes or even different signs of the coefficient on income inequality.

To be able to control for these possible differences interaction terms are introduced into the model. For the purpose to control for the level of democracy, the Polity IV variable is used which ranges from  $-10$  for an autocratic regime to  $+10$  for a democratic regime. An interaction term with the change in income inequality and the democracy variable is created. As reported in *Table 6* the coefficient for this variable is not significant.

To further examine this issue a dummy variable for democracies and non-democracies is introduced. It is recommended by the researchers behind the Polity IV variable that when using it as an indicator of regime categories, it is appropriate to label countries between  $+6$  and  $+10$  as democracies. The coefficient for the interaction term between the democracy dummy and change in income inequality is not significant either. Lastly, the dataset is divided in two groups after the democracy dummy. Even as neither of the two coefficients are statistically different from zero, a difference in the magnitude of the coefficient can be observed, where the coefficient for democracy countries are larger.

The same method is applied when controlling for differences in the coefficient on income inequality among high and low income countries. The coefficient for the interaction term on the income level and the change in income inequality is not significant. When constructing the dummy variable for high income countries 1975, the levels from WDI are used (up to \$3,945 for low and lower middle income countries and from \$3.945 for higher middle and high income countries), however WDI uses GNI (Gross national income) per capita and since no data could be found for this variable in 1975, GDP per capita is used in this study. This coefficient is not either significant. When the sample is divided between high

and low income countries a difference in the magnitude of the coefficient can be observed, however neither of them are statistical significant.

When looking at the results from the regressions done on a divided sample, it should be taken in to consideration the small number of observations. Especially critical is the regression over low income countries as it only includes 21 countries.

*Table 6. Regression result when controlling for development and democracy level.*

$\Delta$ Education Gini	(1)	(2)	(3)	(4)	(5)	(6)
Democracy level* $\Delta$ income Gini	-0.001					
	[0.016]					
Democracy dummy * $\Delta$ income Gini		0.002				
		[0.255]				
Democracy (divided sample)			0.156			
			[0.155]			
Non-democracy (divided sample)			0.030			
			[0.187]			
Income level* $\Delta$ income Gini				0.029		
				[0.130]		
High income dummy* $\Delta$ income Gini					0.133	
					[0.26]	
High income (divided sample)						0.153
						[0.138]
Low income (divided sample)						0.009
						[0.233]

\*,\*\* and \*\*\* denotes statistical significance at 10%, 5% and 1% levels, respectively.

Standard errors in brackets.

Following Gutierrez-Sourdis, this study also controls for age structure affecting inequality in education. The fraction of the population under age 15 is used as a proxy for number of children per household. This variable is included into the model as societies with many children per household tend to be more unequal. Likewise, the fraction of the population over age 65 is controlled for, as societies with a large fraction of elderly tend to be more equal. (Gutierrez-Sourdis, 2006, p. 21) As reported in *Table 7* neither of the two variables are significant at conventional levels and the magnitudes of both coefficients are small. Hence, population structure does not seem to affect inequality in education.

This study also controls the result for differences among regions. Due to the number of observation in the dataset being limited to 64, only three regions are controlled for. The decision over what regions to control for is based on the ten countries with the largest change in education inequality. South Asia, Sub-Saharan Africa and Middle East and North Africa are the three regions with most countries represented. Whereas South Asia and Sub-Saharan Africa show insignificant result, the coefficient on Middle East and North Africa is significant at a 10% level. This means that the coefficient on income inequality for Middle East and North Africa is significantly different from the regions excluded in the regression (Europe and Central Asia, East Asia and Pacific, Latin America and the Caribbean and North America). Middle East and North Africa hence is the only group of countries where a relationship between income inequality and education inequality is significant at conventional levels in the full model (column six in *Table 5*). The coefficient for Middle East and North Africa are positive and of a larger magnitude than in the baseline model. Interesting to notice is that the coefficient on Sub-Saharan Africa is negative however not significant.

*Table 7. Regression result when additional control variables included.*

$\Delta$ Education Gini	(1)	(2)
Population < age 14	-0.000	
	[0.002]	
Population > age 65	-0.003	
	[0.004]	
SA* $\Delta$ income Gini		0.766
		[0.510]
SSA* $\Delta$ income Gini		-0.224
		[0.278]
ME and NA* $\Delta$ income Gini		0.867*
		[0.456]

\*,\*\* and \*\*\* denotes statistical significance at 10%, 5% and 1% levels, respectively.

Standard errors in brackets.

The result from the baseline regression is also controlled for outliers driving the result. When controlling for this issue, observations lying two standard deviations or more away from the mean, in change in education Gini coefficient, are excluded in the regression. This concerns three countries: Arab Republic of Egypt, Islamic Republic of Iran and Uganda. As can be seen from the regression results reported in *Table 8*, no large differences from the baseline

regression result are evident. The coefficient on urban population becomes insignificant while other coefficients remain in both magnitude and significance level.

*Table 8. Regression result from sensitivity analysis, outliers excluded.*

<b>Δ Education Gini</b>	<b>(1)</b>
Δ Income Gini	0.100 [0.119]
GDP per capita	0.022* [0.012]
Δ Human capital	-0.031*** [0.006]
Urban population	-0.000 [0.000]
Education ini 1975	-0.221*** [0.041]
Constant	-0.068 [0.100]
<hr/>	
Number of countries	61
Adjusted R-squared	0.646

\*,\*\* and \*\*\* denotes statistical significance at 10%, 5% and 1% levels, respectively.

Standard errors in brackets.

Another possibly important issue to take into consideration in a regression study is multicollinearity. The correlation matrix is presented in *Appendix 2*. As can be seen, some correlations among variables are rather high, with the highest correlation being at 0.83 between urban population and GDP per capita. However, multicollinearity does not bias the result in a regression but tend to produce large standard errors in the related explanatory variables. This can lead to a false conclusion of no relationship between the dependent and the independent variable. In the baseline regression the coefficient of income inequality is the only insignificant coefficient and hence the only variable for which multicollinearity could be a problem. As previous discussed, it could be suspected that the change in income inequality and initial education inequality would be highly correlated as when the latter variable being introduced into the regression the first variable becomes insignificant. However, the correlation among these two variables is moderate -0.33. In fact, the change in income

inequality is not highly correlated with any of the other independent variables in the model. Therefore multicollinearity is unlikely to be a problem in this study.

Diagnostic testing was performed to test for heteroskedasticity. Breusch-Pagan test was conducted and the null hypothesis of no heteroskedasticity could not be rejected. The possible endogeneity problem has been handled through the model-specification of regressing the change in the independent variable on the change in the dependent variable. However, if endogeneity is still a problem in the model, the coefficients are either downward or upward biased, depending on the miss-specification.

## **7. CONCLUSION**

As become evident in the empirical analysis, in the simple model-specification a positive relation is found between income inequality and education inequality. This relation both decreases in magnitude and significance as the model is extended. However, the coefficient shows a positive relation throughout all six model-specifications. This gives some indication of the hypothesis concerning subsistence constraint and the hypothesis concerning heterogeneity of preferences are more relevant in explaining the investigated relation than the two hypotheses predicting a negative relation.

The coefficients on the control variables included in the model are all statistically significant. Therefore, access to education seems to a larger extent be determined by other variables than income inequality. The positive coefficient on initial GDP shows the interesting result that countries with initial higher GDP per capita worsened their inequality in education. The coefficient is small but still probably reflects the extension of primary education in low income countries and the extension of tertiary education in high income countries during the investigated period. The negative coefficient on urbanization is probably reflecting the constraints on access to education discussed in section 2.1. Children in rural areas are more often constrained by long distance to school, scarcity of qualified teachers and insufficient resources due to the area not being powerful enough to receive political attention.

The statistical insignificance of the coefficient on income inequality might be due to, as the different hypotheses predict, the possible different signs of the relation between the two variables. However, when controlling for differences in both the magnitude and the sign of the relationship, across countries regards to their development and democracy level, no statistically significant differences are found.

This study is to some extent limited due to data restrictions. Regarding the analysis across development and democracy level, the number of observations decreases

under recommended numbers for inference to be valid, in some regressions. With the total dataset covering 64 countries inference for the baseline regression are valid, however if the number of observations could be increased, the number of explanatory variables could as well be increased. Among others, growth rates, poverty and expenditures on education would be interesting variables to include into the regression when investigating access to education. With the possibility to include these variables, further indications about the validity of the hypotheses could be given.

## REFERENCES

- Alesina, Alberto, Reza Baqir and William Easterly. (1999). "Public Goods and Ethnic Divisions." *The Quarterly Journal of Economics* Vol. 114 No. 4 pp. 1243-1284.
- Barro, Robert J. (2000). "Inequality of Growth in a Panel of Countries." *Journal of Economic Growth* Vol.5 pp. 5-32.
- Castelló, Amparo and Doménech, Rafael. (2001). "Human Capital Inequality and Economic Growth: Some Evidence." *Economic Journal*, Vol. 112, No. 478, pp.187-200.
- Checchi, Daniele. (1999). *Inequality in Incomes and Access to Education. A Cross-Country Analysis (1960-90)*. Working Paper Series Department of Economics University of Milan – Bicocca No. 21. To be found at: <http://dipeco.economia.unimib.it/repec/pdf/mibwpaper21.pdf> (2011-05-16)
- Dao, Minh Quang. (2008). "Inequality in Developing Countries: An Empirical Assessment." *Journal of International Finance and Economics* Vol. 8 No. 3 pp.12-18.
- Downs, Anthony. (1957). *An Economic Theory of Democracy* New York: Harper & Row.
- European Group of Research on Equity in Education Systems (2005). *Equity in European Educational Systems A set of indications*. Second edition, (eds. Department of Theoretical and Experimental Education, University of Liège) To be found at: [www.okm.gov.hu/download.php?docID=296](http://www.okm.gov.hu/download.php?docID=296) (2011-05-16)
- Firebaugh, Glenn. (2000). The Trend in Between-Nation Income Inequality. *Annual Review of Sociology* Vol. 26 pp. 323-339.
- Gutierrez-Sourdis, Catalina. (2006). "Inequality and Education Decision in Developing Countries." To be found at: [http://siteresources.worldbank.org/INTDECINEQ/Resources/1149208-1147789289867/IIIWB\\_Conference\\_inequality&education\\_decisions.pdf](http://siteresources.worldbank.org/INTDECINEQ/Resources/1149208-1147789289867/IIIWB_Conference_inequality&education_decisions.pdf) (2011-04-11)
- Gutierrez, Catalina and Ryuichi, Tanaka. (2009), "Inequality and education decisions in developing countries." *Journal of Economic Inequality* Vol. 7 pp. 55-81.
- Heston, Alan, Robert Summers and Bettina Aten. (2009). "Penn World Table". Version 6.3, Centre for International Comparisons of Production, Income and Prices at the University of Pennsylvania, August 2009.
- Holsinger, Donald B. and W. James Jacob (eds.). (2008). *Inequality in Education: Comparative and International Perspectives*. Comparative Education Research Centre, The University of Hong Kong: Springer.
- Hotelling, Harold. (1929). "Stability in Competition." *Economic Journal* Vol. 39 pp. 41-57.
- Marshall Monty G. and Keith Jagers, (Principal Investigators). Ted Robert Gurr (Founder). (2010). "Polity IV Project". Version 2009. Centre for Systematic Peace.

- Mayer, E. Susan. (2001). "How Did the increase in Economic Inequality between 1970 and 1990 Affect Children's Educational Attainment?" *American Journal of Sociology* Vol. 170 No. 1 pp. 1-32.
- Mejia, Daniel and Marc St-Pierre. (2008). "Unequal opportunities and human capital formation". *Journal of Development Economics* Vol. 86 pp. 395-413.
- Meltzer, H. Allan and Scott F. Richard. (1981). "A Rational Theory of the Size of Government." *The Journal of Political Economy* Vol. 89 No. 5 pp. 914-927.
- Roberts, Kevin W.S. "Voting over Income Tax Schedules." *Journal of Public Economics* Vol. 8 pp.329-340.
- Roemer, E. John. (1998). *Equality of Opportunity*. Cambridge: Harvard.
- Solt, Frederick. (2009). "Standardizing the World Income Inequality Database." *Social Science Quarterly* 90(2):231-242. SWIID Version 2.0.
- Sylwester, Kevin. (2002). "Can education expenditures reduce income inequality?" *Economics of Education Review* Vol. 21 pp. 43-52.
- Thomas, Vinod and Yan Wang. (2009). *Distribution of Opportunities Key to Development*. In Holsinger, Donald B. and Wames. J. Jacob (eds.) *Inequality in Education: Comparative and International Perspectives*, Comparative Education Research Centre, The University of Hong Kong: Springer, pp. 34-58.
- Thomas, Vinod, Yan Wang and Xibo Fan. (2001). "Measuring Education Inequality." Policy Research Working Paper 2525, The World Bank.
- Thomas Vinod, Yan Wang and Ramon López. (1998) . "Addressing the Education Puzzle – The distribution of Education and Economic Reform." Policy Research Working Paper 2031, The World Bank.
- United Nations Development Programme. (2010). *Human Development Report 2010 – The Real Wealth of Nations: Pathways to Human Development*. New York: Palgrave Macmillan.
- Wilkinson, Richard G. and Kate E. Pickett. (2005). "Income inequality and population health: A review and explanation of the evidence." *Social Science & Medicine* Vol. 62 pp.1768-1784.
- World Bank, The (2006). *World Development Report 2006: Equity and Development*, New York: Oxford.
- World Bank, The. (2010a). *World Development Indicators*. Washington DC: World Bank.
- World Bank, The. (2010b). *World DataBank, Meta Data*. Washington DC: World Bank.
- World Bank, The. (2011). *World Development Indicators*. Washington DC: World Bank.
- Wu, Kin Bing. (2009). *Education and Inequality in the Developing World*. In Holsinger, Donald B. and W. James Jacob (eds.) *Inequality in Education: Comparative and International*

*Perspectives*, Comparative Education Research Centre, The University of Hong Kong:  
Springer, pp. 86-127.

## APPENDIX

### *Appendix 1. Country list.*

---



---

Argentina ✕	Fiji ✕	Malawi*	Singapore
Australia ✕	Finland ✕	Malaysia	South Africa
Bangladesh* ✕	France ✕	Mauritius ✕	Spain
Belgium ✕	Greece ✕	Mexico	Sri Lanka* ✕
Botswana * ✕	Hungary	Nepal*	Swaziland
Brazil	India* ✕	Netherlands ✕	Sweden ✕
Canada ✕	Indonesia*	New Zealand ✕	Tanzania*
Chile	Iran, Islamic Rep.	Norway ✕	Thailand*
China*	Ireland ✕	Pakistan* ✕	Trinidad & Tobago ✕
Colombia ✕	Israel ✕	Panama	Tunisia*
Costa Rica ✕	Italy ✕	Peru	Turkey
Denmark ✕	Jamaica ✕	Philippines*	Uganda*
Dominican Republic*	Japan ✕	Poland	United Kingdom ✕
Ecuador	Jordan*	Portugal	United States ✕
Egypt, Arab Rep.*	Kenya*	Senegal*	Venezuela, BR ✕
El Salvador	Korea, Rep.	Sierra Leone*	Zambia*

---

\*Low or lower middle income countries 1975 (\$0- \$3,945 GNI per capita)

✕ Democracy countries 1975 (+6 and +10 in Polity IV variable)

### *Appendix 2. Correlation matrix.*

---

	<i>Δ Education</i>	<i>Δ Income</i>	<i>GDP per</i>	<i>Δ Human</i>	<i>Urban</i>	<i>Education</i>
	<i>Gini</i>	<i>Gini</i>	<i>capita</i>	<i>Capital</i>	<i>population</i>	<i>Gini 1975</i>
Δ Education Gini	1					
Δ Income Gini	0.3	1				
GDP per capita	0.51	0.21	1			
Δ Human Capital	-0.47	-0.09	0.03	1		
Urban population	0.4	0.26	0.83	-0.04	1	
Education Gini 1975	-0.76	-0.33	-0.7	0.15	-0.62	1