The Relation Between Income Inequality, Education and Social Structure: Panel Data Analysis of the European Union

Erasmus University Rotterdam

Erasmus School of Economics

Department of Economics

Supervisor: Prof. Dr. J.M.A. Viaene

Name: Melissa Heuchemer

Exam number: 375813

E-mail address: 375813mh@student.eur.nl

Date: 23-06-2016

Abstract

In the last 30 years, income inequality has increased in almost all OECD countries. The purpose of this thesis is to examine the relation between income inequality, education and social structure in the European Union, using the statistics from Eurostat and the OECD. This is researched by using three inequality measures: the Gini coefficient, the Theil index and the Mean Log Deviation. Old age dependency, country of birth and mortality rates are all covered by social structure. The conclusion when using a pooled model is that all these factors do indeed influence income inequality either positive or negative. The fixed effects model concludes that only old age dependency and education influence income inequality. Furthermore, the correlation between the inequality measures is examined. It is concluded that these measures have a very strong correlation.

Table of contents

1. Introduction

Since the 1980s there is a rise in income inequality and in the late 2000s this inequality in Europe was more unequal than in the average OECD country (Fredriksen, 2012). The OECD researched 34 industrial countries in the period 1985-2012 and concluded that the average Gini coefficient increased from 0.29 to 0.32.

In 2012, income inequality measured by the Gini coefficient varied across Europe approximately 10 points (Di Falco, 2014). A lot of factors, such as geographical differences, technology, trade and the role of the government may influence, direct or indirect, income inequality. Income distribution is also an important factor affecting long-run growth (Sarel, 1997).

Research in the United Stated and Mexico shows us that migration can have either a positive or a negative impact on income inequality (Stark et al., 1988). Research in Nicaragua concluded that migration leads to a higher income inequality.

This thesis will try to find a connection between income inequality, education and social structure for the European Union. According to Radcliffe-Brown (1940), social structure is a pattern of relationships between the elements of a society. It can for refer to social groups or individual differences for example. In this case social structure will consist of country of birth, old age dependency and mortality rates.

Since 2013, the European Union consists of 28 countries[1]. The list of EU countries can be found in Appendix A. These 28 countries differ a lot by culture, language, and traditions. Therefore the official slogan of the European Union (EU) is “united in diversity”. Because of all these differences between the countries, using the EU is a perfect way to investigate which factors affect income inequality. The research question is stated as follows:

What is the relationship between income inequality, education, and social structure in the European Union in the period 2006-2014?

To answer the research question, a couple of partial research questions are formulated and will be answered. This is followed by answering the research question.

[1] Source: Rijksoverheid: Rijksoverheid: https://www.rijksoverheid.nl/onderwerpen/europese-unie/vraag-en-antwoord/welke-landen-behoren-tot-de-europese-unie-eu-en-de-europese-economische-ruimte-eer

Jencks et al. (1972) argue that incomes will not be equalized when everyone has the same educational opportunity. While Coleman et al. (1973) argue that this conclusion has several problems like the understatement of educational attainment on income. They conclude that educational attainment is in fact critical for employment and income. This is proven by comparing white and black men in 1968 in the United States (Coleman et al., 1973). Chiswick concludes that income inequality is larger when there is absolute schooling inequality (Chiswick, 1971). De Gregorio and Lee indicate that a higher educational attainment plays a significant role in a more equalized income distribution (De Gregorio and Lee, 2002). Breen and Chung support this and say that the income gap between those with and without a college education is a major cause of increasing inequality (Breen and Chung, 2015). Research also showed that technology and globalization have an impact on income inequality. Nelson and Phelps (1966) concluded that technological developments tend to increase the demand for high-skilled people, as education will improve one's ability to understand and process information (Nelson and Phelps, 1966). Therefore, the first partial research question is:

Does a higher level of education reduce income inequality?

It is expected that higher educational attainment levels will reduce income inequality, as people will have better jobs and earn more money.

The second factor that will be researched in this paper is social structure. First we will research the country of birth or the origin of people. In developing countries, international migration leads to a decrease in the level and depth of poverty and thus decreases income inequality (Adams Jr. and Page, 2005). Borjas predicted that a more unequal income distribution in country of birth will have a negative impact on the skills mix of migrants in the migration country (Liebig and Sousa-Poza, 2004). The question arises if this will cause an increase in the income inequality. Therefore, the second partial research question is:

Does the country of birth affect income inequality?

It is expected when you are born in a foreign country and working in the reporting country, that income inequality will rise. Most foreign workers will do more manual labour and have a lower skill level, which in turn pays less and increases income inequality.

Another factor that is related to the social structure is old age dependency. When people get older and reach the retirement age, they are dependent on their pension for income. Every country has its own social security system or public pension benefits. Research has shown that income inequalities in old age are shaped by retirement income policies (Brown and Prus, 2004). If the retirement age will decrease, poverty will increase and thus income inequality will also increase. After retirement, people have less to spend and will be more dependent on society (Walker, 1980). Therefore, the third partial research question is:

Does a higher old age dependency lead to an increase in income inequality?

It is expected that a higher old age dependency will indeed lead to an increase in income inequality. With more people living of a small retirement than working, income inequality is expected to rise.

The third factor related to the social structure is mortality rates. Research has shown that income inequality is associated with mortality trends (Kaplan et al., 1996). Two indicators, namely the Gini coefficient and the Robin Hood Index, are strongly correlated with mortality in the United States (Kennedy et al., 1996). However, Judge (1995), researched 13 OECD countries and found no association with other inequality measures. Therefore the fourth partial research question is:

Do mortality rates have an impact on income inequality measures?

Mortality rates and income inequality are linked to each other. In countries with high income inequality, people live shorter than in countries with low income inequality. In countries with low income inequality, people tend to live healthier and therefore live longer (Mendelson, 2011). Therefore, it is expected that income inequality will decrease when mortality rates increase.

There are several income inequality measures, where the Gini coefficient is most commonly used. Kennedy et al. (1997), found that six income inequality measures were all highly correlated with each other, with a correlation between 0.86 and 0.99. This means that the choice of income inequality measure doesn't matter (Kennedy et al., 1997). However, Weich et al. (2002), indicated that the choice of an income inequality measure does influence the results. The fifth and last partial research question is:

Do income inequality measures have a very strong correlation with each other?

It is expected that the three income inequality measures have a strong correlation between them. Champernowne (1974) has shown that income inequality measures are correlated between the values 0.802 and 0.996. Where the value of 0.966 is the correlation between the Gini coefficient and the Theil index.

The remainder of this paper is organized as follows. In chapter 2 there will be a review of previous researches on income inequality and which factors influence this. In chapter 3 some theory will be explained. Chapter 4 focuses on the data and methodology used. The results will be shown in chapter 5. Chapter 6 contains the conclusion and chapter 7 will contain a discussion and some recommendations for future research.

2. Literature Review

When a country is experiencing economic growth, will income inequality increase or decrease? This is the question that Simon Kuznets (1955), one of the first to research income inequality, wanted to answer. The results show that in time of industrialization income inequality will increase, but when industrialization continues the income inequality will decrease (Kuznets, 1955). This effect is called the Kuznets curve later on.

Rodgers (1979) investigated the effects of income and inequality on mortality. This was examined using an international cross-section analysis. He concluded that the results for life expectancy at birth suggests a difference in average life expectancy between relative equal countries and relative unequal countries. This difference in life expectancy can be as much as five to ten years (Rodgers, 1979). Kawachi et al. (1997) took it a bit further and chose to investigate the relationship between social capital, income inequality and mortality. The conclusion supports the research of Rodgers in 1997, that income inequality leads to increased mortality, but they also say that disinvestment in social capital is a key factor that will increase mortality (Kawachi et al., 1997).

Muller (2002) argues that looking at mortality and income inequality isn't enough and added education to the model, as education precedes working and thus income. When adding education, the fit of the regression improved significantly. He concluded that lack of high school education accounts for income inequality and is a powerful predictor of mortality rates in the United States (Muller, 2002).

Kennedy et al. (1998) researched the effects of social capital on income inequality. They concluded that income and poverty are predictors of crime. Also, the gap between the rich and the poor grows through the undermining of social capital. This decreasing level of social capital creates violent crime. One example of increased violent crime is an increasing level of firearm homicide (Kennedy et al., 1998).

There also exists a relationship between income inequality and health. In this relationship, income distribution is intended to measure the scale of social class differences in a society. This only holds for large countries, as in small countries it is unlikely that income inequality reflects the classification of the population into social groups (Wilkinson and Pickett, 2006).

Oshio and Kobayashi (2010) researched the relationship between income inequality and self-rated health in Japan. Individuals who live in areas with high inequality reported that they

are unhappy and unhealthy, even after controlling for individual characteristics as gender and educational attainment (Oshio and Kobayashi, 2010).

Another research shows that income inequality and poverty are related to policies that reduce corruption. This is important because corruption distorts the role of the government and it is the government that allocates resources. When there is less corruption, the government can allocate resources better and there is less income inequality (Gupta et al., 2002).

When looking at the relationship between remittance income, migration and income inequality, it is concluded when remittances are a substitute for home earnings that income inequality is higher in regions with migrant households (Barham and Boucher, 1998).

As you can see, a lot of research in the field income inequality has been done. What all these researches show, is that they all denote different sources for income inequality. This paper will mainly focus on education and the social structure of a country and at the end of the paper it will be clear whether these factors have significant influence.

3. Theoretical Framework

Before the relationship between income inequality, education and social structure will be examined, there will be a more detailed explanation of the theories and concepts.

3.1. Inequality Measures

Income equality means that everyone in a society will have the same income. However, this is only possible in theory. In practice we see that there is actually income inequality. Income inequality can be measured through the Gini coefficient, Hoover index, Theil index, Mean Logarithmic Deviation (MLD) or the Atkinson index (Sala-i-Martin, 2002).

According to Ray (1998) there are four criteria for inequality measurement. The first criterion is anonymity. It doesn't matter if person A earns X and person B earns Y. The only thing that matters is the distribution of income; X and Y should be viewed as identical. The second criterion is the population principle. This principle states that, if we replicate the entire population, income inequality will not change and is the same in the two distributions. It tells us that the size of a country doesn't matter, the proportions of who earns what are important. The third criterion is the relative income principle. This principle states that only relative incomes matter and absolute incomes should not. The last criterion is the Dalton principle. This principle states that if a rich person receives income from a poor person, income inequality will increase. But when a poor person receives from a rich person, income inequality will decrease. This is called a regressive transfer (Ray, 1998).

Gini Coefficient

The most commonly used method is the Gini coefficient, which measures the difference between wage and the percentage of workers in a country. The Gini coefficient can measured using the Lorenz curve. The Lorenz curve is drawn in Figure 3.1. The figure shows cumulative percentages of the population as if the population is sorted in increasing order of incomes. At point X, approximately 40% of the population earns 10% on the income axis. This means that the poorest 40% earn 10 % of the overall income. At point Y, approximately 90% of the population earns 80% of the overall income. If everyone earns the same income, the line will lie on the 45° degree line, where there is perfect equality. This means that the poorest 20% will earn exactly 20% of the overall income. When income inequality increases, the Lorenz curve will fall further below the perfect equality line. Inequality is thus the area between the perfect equality line and the Lorenz curve; area A (Ray, 1998).

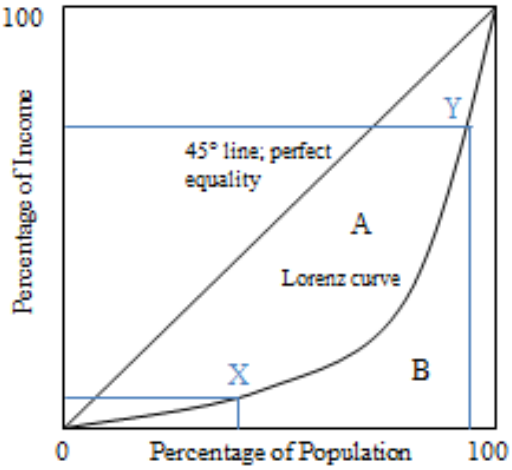

Figure 3.1: Lorenz Curve.

On the Y-axis, the percentage of Income and on the X-axis, the percentage of Population is found.

Source: based on the Lorenz curve of Gastwirth (1972).

The Gini coefficient is the ratio of the area between the Lorenz curve and the 45° line; area A. The index can take a value anywhere from 0 to 1. When a society is completely equal, the value will be 0, and when a society is completely unequal the value will be 1 (Atkinson and Bourguignon, 2000).

The Gini after social transfers is most commonly used (Atkinson, 1997). The equation for the Gini coefficient is derived from Eurostat and is calculated as follows:

$$(3.1)\quad Gini = \left(\frac{2 \times \sum_{i=first\ person}^{last\ person} \left(A \times E \times \sum_{i=first\ person}^{person\ i} A \right) - \sum_{i=fist\ person}^{last\ person} (A^2 \times E)}{\sum_{i=first\ person}^{last\ person} (A) \times \sum_{i=first\ person}^{last\ person} (A \times E)} \right) - 1 \times 100$$

where

G is the Gini coefficient

A is the adjusted cross sectional weight for person i

E is the equivalised disposable income after social transfers

i is a person, where persons are sorted from lowest to highest value

The Gini coefficient is based on the equivalised disposable income of each individual. The equivalised disposable income is calculated using the modified OECD equivalent scale. This means that this scale gives a weight of 1.0 to a first adult, 0.5 to a second adult (or a person aged 14 and over) and 0.3 to children under the age of 14[2].

[2] Source: Eurostat : http://ec.europa.eu/eurostat/statistics-explained/index.php

The advantage of using the Gini coefficient is that it measures inequality by means of a ratio, rather than a variable that doesn't represent the entire population. The Gini coefficient can therefore be used to compare income distributions across different countries. The Gini coefficient satisfies all four criterion mentioned above. With the Gini coefficient it doesn't matter who earns a high income and who earns a low income. It also doesn't consider the size of the economy or the size of the population. The Gini coefficient does satisfy the Dalton principle, also known as the transfer principle; if income transfers from a rich person to a poor person, income inequality and thus the Gini coefficient will decrease.

Hoover Index

The Hoover index, called after Edgar Hoover (1936) is also called the Robin Hood index, because it describes how much income has to be transferred from the rich to the poor in order to reach an equal distribution. The Hoover index is similar to the Gini coefficient in the sense that it also takes a value between 0 and 1, with 0 being totally equal. The Hoover index can also be derived from the Lorenz curve in Figure 3.1. The Hoover index is the largest difference between the 45° line of total equality and the Lorenz curve (Hoover, 1936).

Theil Index

The Theil index is based on statistical information theory. It has a large sensitivity to income transfers from the poor to the rich. The Theil line will increase in steepness as the transfers grow. The Theil index takes individual characteristics into account. Therefore, the Theil index is preferred when there are inequalities in subgroups, for example in income groups. The Theil index is not commonly used because of its complexity (Conceição and Ferreira, 2000). Theil's T, also called General Entropy, measure may be calculated using the following equation:

$$GE(1) = \sum_{i=1}^{n} p_i \ln(\frac{p_i}{q_i}) \tag{3.2}$$

where

p_i is the income share of group i

q_i is the population share of group i

i is person i, persons are sorted from lowest to highest value

One of the conditions to calculate the index is that incomes are non-negative. The value of the Theil index lies between 0 and 1, where 0 is an equal distribution (Conceição and Ferreira, 2000).

Mean Log Deviation

Mean Log Deviation, further referred as MLD, is another Generalized Entropy measure with coefficient 0 and is also called Theil's L index (Sala-i-Martin, 2002). The measure is calculated using the following equation:

$$GE(0) = \sum_{i=1}^{n} q_i \ln(\frac{q_i}{p_i}) \tag{3.3}$$

where

p_i is the income share of group i

q_i is the population share of group i

i is person i, persons are sorted from lowest to highest value

The value of the MLD lies between 0 and 1, where 0 means an equal distribution and 1 means an unequal distribution (Sala-i-Martin, 2002)..

Atkinson Index

The Atkinson index takes judgments about social welfare into account. The index is derived by calculating the equity-sensitive average income. This average income is defined as that level of per capital income which would make total welfare equal to the total welfare by the actual income distribution. The index also takes a value between 0 and 1 (Kawachi and Kennedy, 1997).

3.2. Education

According to the human capital model of income distribution by Schultz (1961) and Becker (1964), the distribution of earnings is determined by the level and the distribution of schooling across the population. The model says that there is a positive association between educational inequality and income inequality. Educational inequality is measured by the variance of schooling. Furthermore, the model predicts that the effect of increased average schooling on income inequality may either be negative or positive. The effect depends on the evolution of the rates of return on education (Blaug, 1976).

3.3. Social Structure

Social structure is the pattern of relationships between the elements of a society, according to Radcliffe-Brown. The relationships are built up by the persons living in a society. Social structure is different from social relations, which is a network of social relations involving other persons. Social structure can refer to social groups, such as nations and tribes, but it can also refer to individual differences, such as sex (Radcliffe-Brown, 1940).

Another, more recent, topic on inequality is horizontal inequality. According to Krugman (2016), horizontal inequality is the term of art for inequality measured, not between individuals, but between racially or culturally defined groups. When enough data is available on average group performance, income inequality measures can indicate the dispersion of achievements among groups (Steward, 2000). Religion, citizenship, racial, social clubs and age are some examples of horizontal groups (Steward et al., 2005). Religion statistics or statistics on social clubs and races are not available. Being a member of a group means that being a group member leads to different treatment by others and that being a member feels like it constitutes a significant aspect of your identity (Steward et al., 2005). This makes it hard to see who belongs to which group, because it is often subjective. Data on inequalities of specific groups are not or scarcely available thus are not included. An aspect of citizenship is included: the country of birth. Old age dependency and mortality rates are both aspects of age.

In this paper, social structure will consist of country of birth, old age dependency and mortality rates.

Country of Birth

The country of birth is taken into account to see whether it has significant effect on income inequality. Country of birth can be divided into three categories, namely: born in reporting country, born in the EU except reporting country and born in foreign country. The measurement of people born in the EU that reside in the reporting country is often difficult. This is because people in the EU are free to move from one country to another. Also, the EU has grown a lot in the last couple of years, so the measurements are incomplete or biased. Birth in the EU except reporting country is thus not included. According to Eurostat, when you are born in any other country than the reporting country, you fall under the category birth in foreign country. This covers people who have migrated from their country of birth to their current country of residence. The countries which people originated from are not included, because there is not enough data available on this subject.

Old Age Dependency

Old age dependency is the ratio between the number of persons aged 65 and over and the number of persons aged between 15 and 64[3].

Mortality Rates

Mortality rate, or also known as the death rate, is the number of deaths in a population. There are several categories; mortality at birth, life expectancy at 65, mortality rate, infant mortality rate, and maternal mortality rate. We will only look at the life expectancy at birth. This is the most commonly used indicator and is the number of years that a person can expect to live at birth[4].

[3] Source: Eurostat: Old-age-dependency ratio: http://ec.europa.eu/eurostat/tgm/table.do?tab=table&init=1&language=en&pcode=tsdde510&plugin=1

[4] Source: Eurostat: Mortality and life expectancy statistics: http://ec.europa.eu/eurostat/statistics-explained/index.php/Mortality_and_life_expectancy_statistics

4. Data and Methodology

In this chapter, the data and methodology will be discussed. The investigated period will be 2006-2014. The data is retrieved from the statistical office of the European Union, Eurostat, or the OECD. Data for all countries before 2006 is not available or is incomplete.

4.1. Data

Panel data is used in this thesis, as we have both time series (2006-2014) and cross-sectional data (28 EU countries) (Brooks, 2014). The use of panel data also ensures control for unobserved individual-specific characteristics. These characteristics are included in the error term of the regression (Carter Hill et al., 2012).

The investigated period will be 2006-2014. The data will be retrieved from the statistical office of the European Union, Eurostat, or the OECD. Data on income inequality, GDP per capita, educational attainment level, statistics on country of birth, old age dependency and mortality rates are needed. The data used can be found in Appendix B.

The Dependent Variables

Gini Coefficient *(GINI)*

One of the dependent variables is the Gini coefficient. The Gini coefficient for disposable income is calculated using Equation 3.1. The data is retrieved from Eurostat.

Theil Index (*Theil)*

The Theil index is the second dependent variable. The Theil index is calculated using Equation 3.2. The data on income shares and population shares is retrieved from Eurostat.

Mean Logarithmic Deviation *(MLD)*

The MLD is the third and last dependent variable. The MLD is calculated using Equation 3.3. The data on income shares and population shares is retrieved from Eurostat.

The Independent Variables

The independent or control variables used are GDP, educational attainment country of birth, old age dependency and mortality rates. These variables will be described below. Control

variables are included, because when they are left out, an omitted variable bias can arise (Carter Hill et al., 2012).

GDP *(GDP)*

GDP is calculated by adding up the consumption, investment, and government spending and subtracting net exports Equation 4.1.

(4.1) $$GDP = C + I + G + (X - M)$$

Another way to calculate GDP is also called the income approach, as seen in Equation 4.2.

(4.2) $$GDP = wages + self\ employment\ income + rent + interest + profits + indirect\ business\ taxes + depreciation + net\ income\ of\ foreigners$$

Real GDP per capita is used. Previous research revealed that using real GDP leads to better results when included in the Gini coefficient equation (Barro, 2000). Real GDP measures the average real income in a specific country. Real GDP does not include negative effects of economic activity. Not all countries of the EU have the euro as currency; this makes a comparison a little more complicated. Therefore, real GDP in euros is used. The data on real GDP per capita is retrieved from Eurostat.

Year *(Year)*

The years 2006-2014 are used. To control for year effects, 2006 is used as reference year.

Country *(Country)*

The data for 28 countries is included. These are the 28 countries constituting the EU. The list of the EU countries can be found in Appendix A.

Educational Attainment Level *(Educ_att_x)*

Scholar systems are different in all European countries; therefore we need a common criterion. Lopez et al. were one of the first to split education into attainment levels to see the effects of a higher education on the Gini coefficient. The education attainment levels were divided into three categories: primary, secondary and tertiary education (Thomas et al., 2001). The categories can be seen in Table 4.1. The data is in percentages of the population aged 15-65. The data is retrieved from Eurostat.

Level	Type of education	Abbreviation
0-2	Less than primary, primary and lower secundary	*(Educ_att_low)*
3-4	Upper secondary and post-secondary non-tertiary	*(Educ_att_med)*
5-8	Tertiary	*(Educ_att_high)*

Table 4:1: Education Attainment Level

As the educational attainment level rises, it is expected that income inequality and thus the Gini coefficient, the Theil index and the MLD decrease. When more people have a higher level of education and less people have a low educational level, people will have better jobs and earn more. When more people have higher incomes and less people have lower incomes, income inequality will decline.

Country of Birth *(X_country)*

Parentage or origin is the second independent variable. The country of birth will thus be divided into two categories as seen in Table 4.2. The data is in percentages of total population. Only the variable Foreign will be added into the equation. The variables Foreign country and Reporting country are perfectly negatively correlated, because you either belong to the Foreign category or the Reporting category. Only *Foreign_country* is added because we want to see if there is a relationship between origin and income inequality. The foreign-born population data shown here includes people born abroad as nationals of their current country of residence.

Birth country	Abbreviation
Reporting	*(Reporting_country)*
Foreign	*(Foreign_country)*

Table 4:2 Country of Birth

When you are born in a foreign country and you migrate to another country it is common that you do this for job opportunities. The Netherlands for example has a lot of guest or seasonal workers. A lot of these workers perform tasks in greenhouses, work long hours and earn a low wage. They do jobs that the Dutch won't because of working conditions. The idea behind this, is that the more people that originated from other countries that work in the reporting country, the higher income inequality will be because of the low wages paid to seasonal workers. The data is retrieved from Eurostat.

Old Age Dependency *(Oldagedependency)*

The old age dependency ratio is the ratio between the number of persons aged 65 and over and the number of persons aged between 15 and 64 years old. The value is expressed as the number of persons age 65 and over per 100 persons of the working age, 15-64. 65 and over is chosen because this is the average age that persons become economically inactive or reach the pension age. The age group 15-64 is chosen, because you are able to live independently from age 15-16. The old age dependency in Belgium for example is 27.3 in 2014. This means that there are on average 27.3 persons aged 65 and over on 100 persons aged 15-64. The data is retrieved from Eurostat.

When the old age dependency ratio grows, there are more "old" people then "young" people. The "old" people do not work anymore and they must survive on their pension. When those pensions are low, they do not get to spend their money on extras and barely scrape by. "Young" people do have jobs and earn money. Therefore, it is expected when there are more "old" people in a society, income inequality will increase, by the growing gap between the rich (young people) and the poor (old people).

Mortality Rates *(Mortalityrates)*

Mortality rates is the life expectancy at birth. This means if it takes a value of 80.7 in Belgium, you are expected to live for 80.7 years when you are born. The data is retrieved from Eurostat and consists of the total population; men and women are not separated.

4.2. Methodology

There are three methods of estimating the data: pooled, random effects, or fixed effects. In a pooled model data of different individuals are pooled together without inclusion of individual differences. This means that the coefficients are all the same. The random effects model also doesn't include individual differences (Brooks, 2014). The difference between the random effects model and the fixed effects model is that the fixed effects model allows for endogeneity and individual differences (Mundalk, 1978). Prior research on income inequality compared the pooled model with the fixed and the random model (Lorgelly and Lindley, 2008). The pooled model will be compared with the fixed effects model or the random effects model. A Hausman test will be conducted to see which of the two models will be used. The Hausman test compares the coefficients from the random model to those of the fixed effects model. In large samples, the random effects and the fixed effects estimates should be similar

(Carter Hill, 2012). If the null hypothesis that the estimators yield identical results is rejected, the fixed effects model must be used (Hausman, 1978). Data analysis will be done using Stata in combination with excel (data collection, arranging and normalization).

Structural Form

The following equations will be used for the estimation of the inequality measures. First the GINI coefficient will be estimated using Estimation (4.3). The equations for the Theil index and the MLD are the same, only the coefficients differ. The variable *Country* is not included in the random and fixed effects model as *Country* is the panel variable and *Year* is the time variable. For the pooled model, the variable *Country* is included.

(4.3)

$$GINI_{it} = \beta_0 + \beta_1 GDP_{it} + \beta_2 Country_{it} + \beta_3 Year_{it} + \beta_4 Educ\_att\_low_{it} + \beta_5 Educ\_att\_med_{it} + \beta_6 Educ\_att\_high_{it} + \beta_7 Foreign\_country_{it} + \beta_8 Oldagedependency_{it} + \beta_9 Mortalityrates_{it} + \varepsilon_{it}$$

The Theil index will be estimated using Estimation (4.4).

(4.4)

$$Theil_{it} = \beta_0 + \beta_1 GDP_{it} + \beta_2 Country_{it} + \beta_3 Year_{it} + \beta_4 Educ\_att\_low_{it} + \beta_5 Educ\_att\_med_{it} + \beta_6 Educ\_att\_high_{it} + \beta_7 Foreign\_country_{it} + \beta_8 Oldagedependency_{it} + \beta_9 Mortalityrates_{it} + \varepsilon_{it}$$

The Mean Log Deviation will be estimated using Estimation (4.5).

(4.5)

$$MLD_{it} = \beta_0 + \beta_1 GDP_{it} + \beta_2 Country_{it} + \beta_3 Year_{it} + \beta_4 Educ\_att\_low_{it} + \beta_5 Educ\_att\_med_{it} + \beta_6 Educ\_att\_high_{it} + \beta_7 Foreign\_country_{it} + \beta_8 Oldagedependency_{it} + \beta_9 Mortalityrates_{it} + \varepsilon_{it}$$

Endogeneity

The endogeneity problem can occur for three reasons. The first reason is omitted variables bias. This means that there is not enough data available. The second reason is measurement error. This means that it isn't possible to get a perfect measure of the independent variable(s). The third reason is simultaneity. This means that some independent variables are included in the dependent variable. The independent variables will be correlated with the error term. It is also possible that two or more variables simultaneously affect each other (Antonakis, 2014).

Some variables in the Equations 4.3, 4.4 and 4.5 can influence each other. Prior research has shown that *Education* had a causal impact on *Mortality rates*. It was concluded that the more-educated population live longer (Lleras-Muney, 2005). The variables *Mortality rates* and *Country of birth* can influence each other. Previous research has shown that geographical patterns can influence death rates. The health status of migrants differs considerably from non-migrants. The difference in health status grows considerably by the distance migrants travel to reach their end destination. The further they have to travel, the bigger the gap in health status. In countries where there is net out-migration, people tend to be less healthy, and this in turn influences the *Mortality rates* (Bentham, 1988). The variables *Education* and C*ountry of birth* can also affect each other. Prior research has shown that migration has a significant negative impact on education attainment level (McKenzie and Rapoport, 2006).

Assumptions Linked to Least Squares Methods

The estimation of the models will be done using the Least Squares method. To use this method, five assumptions must be met.

(I)	$E(u) = 0$	expectation of the residual is zero
(II)	$E(u^2) = 0$	no heteroscedasticity
(III)	$Cov(u_i, u_j) = 0$	no serial correlation
(IV)	$Cov(u_t, u_x) = 0$	endogeneity, no errors in variables
(V)	$u_t \sim N(0, \sigma^2)$	disturbances are normally distributed

If the assumptions I-IV hold, the estimators determined by Ordinary Least Squares (OLS) will be known as best linear unbiased estimators (BLUE). If a number of assumptions are met, OLS is unbiased, efficient and consistent (Brooks, 2014). For the pooled model, the assumptions will be tested. The fixed effect model corrects violations (III) and (IV) itself. When using a pooled model, these violations are not corrected automatically.

If a constant term is included in the equation, assumption (I) is met. Assumption (II) can be tested using the Breusch-Pagan test for heteroskedasticity. To test for serial correlation (III), a Durbin Watson test can be conducted. For assumption (IV) an OVtest can be conducted, this is a Ramsay regression specification-error test for omitted variables. For assumption (V) a Jarque-Bera test can be conducted on residuals (Brooks, 2014).

When adding the robust option in Stata, there is no need to test for heteroskedasticity, normality and serial correlation. With the robust option, the standard errors deal with minor concerns about these assumptions and corrects automatically if these problems arise (Williams, 2015). A constant term is also included in the model; this means that only assumption IV needs to be tested; a test for endogeneity (Carter Hill, 2012).

5. Results

To answer the research question, the pooled model will be tested first. The fixed effects model is tested second followed by the random effects model. The models are tested for all three inequality measures. For the *GINI* coefficient; Equation (4.3), the *Theil* index; Equation (4.4), and for the *MLD*; Equation (4.5) is used. The results will be discussed first for the *GINI* coefficient, second for the *Theil* index and followed by the results of the *MLD*.

5.1. Pooled Model

The assumptions (II)-(IV) need to be tested first. Assumption (V) was tested first. The results are shown in Appendix C for all three inequality measures. It is obvious that assumption (V) is not met, Kurtosis is far removed from 3 and Skewness is far removed from 0. Therefore, the option robust is used for all inequality measures, which corrects automatically for assumptions (II), (III) and (V). Assumption (IV) is the only assumption that needs to be tested; this is done using a Ramsey Reset test.

GINI

Assumption (IV) is tested first. The test result is shown in Table 5.1. The P-value is higher than 0.05, the null hypothesis of no omitted variables cannot be rejected. This means that all of the five assumptions are met.

	GINI
P-value	0.1687

Table 5.1: Ramsey RESET test GINI

Table 5.2, the second column (1) to the left shows us the results if the *GINI* coefficient is used as the dependent variable. The first value is the coefficient and the value between the brackets is the p-value. *GDP* and *Mortalityrates* are significant at a 5% level and have a negative impact on the *GINI* coefficient. This means that a higher GDP and a higher mortality rate decrease the *GINI* coefficient, and thus lowers income inequality. *Foreign_country* and *Oldagedependency* are also significant at a 5% level and have a positive impact on the *GINI* coefficient. This means that people born in a foreign country and a higher old age dependency increase the *GINI* coefficient. The educational levels are not significant at a 5% level, but are significant at a 10% level. When they are significant, they decrease the *GINI* coefficient and thus decrease the level of income inequality. The variables *Country* and *Year* are insignificant at a 5% and 10% level.

Table 5.2: Pooled model with three educational levels

VARIABLES	(1) GINI	(2) Theil	(3) MLD
GDP	-1.21e-06***	-1.16e-06***	-1.14e-06***
	(1.52e-09)	(2.92e-10)	(1.75e-08)
Country	0.000699	0.000492*	0.000648**
	(0.112)	(0.0879)	(0.0331)
2007.Year	0.00751	0.00250	0.00299
	(0.691)	(0.829)	(0.803)
2008.Year	0.00796	0.00304	0.00351
	(0.670)	(0.790)	(0.766)
2009.Year	0.00333	-0.000735	0.00112
	(0.856)	(0.946)	(0.922)
2010.Year	0.0129	0.00280	0.00622
	(0.372)	(0.768)	(0.532)
2011.Year	0.0169	0.00585	0.00956
	(0.269)	(0.548)	(0.357)
2012.Year	0.0145	0.00362	0.00664
	(0.334)	(0.708)	(0.519)
2013.Year	0.0137	0.00443	0.00786
	(0.335)	(0.644)	(0.441)
2014.Year	0.0170	0.00854	0.0124
	(0.220)	(0.373)	(0.236)
Educ_att_low	-9.481*	-7.167*	-6.941*
	(0.0824)	(0.0620)	(0.0975)
Educ_att_med	-9.661*	-7.330*	-7.119*
	(0.0768)	(0.0564)	(0.0893)
Educ_att_high	-9.413*	-7.133*	-6.915*
	(0.0843)	(0.0634)	(0.0989)
Foreign_country	0.188***	0.164***	0.166***
	(1.13e-07)	(2.56e-08)	(3.36e-07)
Oldagedependency	0.222***	0.246***	0.321***
	(0.000508)	(2.38e-06)	(6.33e-09)
Mortalityrates	-0.638***	-0.650***	-0.770***
	(2.37e-06)	(4.47e-10)	(0)
Constant	10.28*	7.829**	7.688*
	(0.0581)	(0.0413)	(0.0662)
Observations	252	252	252
R-squared	0.262	0.401	0.422

Notes: (i) Output for GINI (1), Theil (2) and MLD (3). The first value is the coefficient and the value between the brackets is the p-value. (ii) Robust pval in parentheses *** p<0.01, ** p<0.05, * p<0.1.

Theil

Assumption (IV) is tested first. The test is shown in Table 5.3. The P-value is higher than 0.05, the null hypothesis of no omitted variables cannot be rejected. This means that all of the five assumptions are met.

	Theil
P-value	0.0840

Table 5.3: Ramsey RESET test Theil

Table 5.2, the third column (2) to the left shows us the results if the Theil index is used as the dependent variable. *GDP* and *Mortalityrates* are significant at a 5% level and lower the *Theil* coefficient; decreasing income inequality. *Foreign_Country* and *Oldagedependency* are also significant at a 5% level increase the *Theil* index. The educational levels and the variable *country* are not significant at a 5% level, but are significant at a 10% level. The educational levels, when significant, deceases the *Theil* index. The variable *Country*, when significant, has a positive impact on the *Theil* index. The variable *Year* is insignificant at a 5% and 10% level.

MLD

Assumption (IV) is tested first. The test is shown in Table 5.4. The P-value is lower than 0.05, the null hypothesis of no omitted variables cannot be rejected. This means that four of the five assumptions are met.

	MLD
P-value	0.0363

Table 5.4: Ramsey RESET test MLD

Table 5.2, the fourth and last column (3) to the left shows us the results if the *MLD* is used as the dependent variable. *GDP* and *Mortalityrates* are significant at a 5% level and have a negative relation with the *MLD*, decreasing the income inequality. *Foreign_Country*, *Country* and *Oldagedependency* are also significant at a 5% level and have a positive relation with the *MLD*, increasing the income inequality. The educational levels are not significant at a 5% level, but are significant at a 10% level. When significant, the educational levels have a negative relation with the *MLD*, decreasing income inequality. The variable *Year* is insignificant at a 5% and 10% level.

The three Educational levels are not significant at a 5% level for all three inequality measures. Willis and Rosen (1987), considered only two levels of schooling. These were high school and higher than high school. Barro and Lee (2013) also divided the educational levels into two broad categories. The first category is the less-educated population, this consists of uneducated people and people who have reached the primary level of schooling. The second category is the more-educated population and consists of people who have reached at least the secondary level of schooling.

When adding the variables *Educ_att_med* and *Educ_att_high* and call it *Educ_att_medhigh*, only two variables of educational attainment are created; *Educ_att_low,* the less-educated population and *Educ_att_medhigh,* the more educated population. These two variables are perfect negatively correlated with each other, thus only one of these two variables is included in the model. To see whether a higher education has an impact on income inequality, only the variable *Educ_att_medhigh* is included for education in Equation 4.3, 4.4 and 4.5. The results are shown in Table 5.5.

The results differ slightly from the other model. The variable *Education_medhigh* is now significant even at the 5% level for all three inequality measures. This means a higher education lowers income inequality. The variable *Country* is only significant at a 10% level for the *Theil* index and significant at a 5% level for the *MLD.* For the variables *Foreign_country* and *Oldagedependency* nothing has changed; they are still significant at a 5% level and increase the inequality measures. For the variables *GDP* and *Mortalityrates* also nothing has changed, they are still significant at a 5% level and decrease the inequality measures. The variable *Year* has changed however, the year 2014 is significant at a 5% level for the *MLD* and is significant at a 10% level for the *GINI* coefficient and *Theil* index. The year 2014 has a positive relationship on the inequality measures and thus increases income inequality.

Table 5.5: Pooled model with one educational level

VARIABLES	(1) GINI	(2) Theil	(3) MLD
GDP	-9.67e-07***	-9.61e-07***	-9.40e-07***
	(1.25e-05)	(1.14e-06)	(1.99e-05)
Country	0.000718	0.000506	0.000663**
	(0.120)	(0.102)	(0.0414)
2007.Year	0.0110	0.00517	0.00561
	(0.596)	(0.685)	(0.673)
2008.Year	0.0118	0.00599	0.00642
	(0.562)	(0.628)	(0.615)
2009.Year	0.00826	0.00308	0.00493
	(0.679)	(0.796)	(0.695)
2010.Year	0.0197	0.00802	0.0114
	(0.240)	(0.456)	(0.313)
2011.Year	0.0219	0.00977	0.0136
	(0.184)	(0.351)	(0.221)
2012.Year	0.0221	0.00960	0.0128
	(0.179)	(0.357)	(0.247)
2013.Year	0.0248	0.0131	0.0166
	(0.130)	(0.218)	(0.138)
2014.Year	0.0312*	0.0196*	0.0236**
	(0.0576)	(0.0702)	(0.0404)
Educ_att_medhigh	-0.130***	-0.125***	-0.138***
	(3.71e-08)	(5.35e-09)	(8.26e-10)
Foreign_country	0.238***	0.205***	0.208***
	(6.99e-09)	(2.84e-10)	(4.10e-09)
Oldagedependency	0.192***	0.222***	0.296***
	(0.00646)	(0.000152)	(1.51e-06)
Mortalityrates	-0.501***	-0.541***	-0.657***
	(0.00109)	(4.09e-06)	(1.48e-07)
Constant	0.704***	0.586***	0.669***
	(3.00e-07)	(9.34e-09)	(7.68e-10)
Observations	252	252	252
R-squared	0.192	0.321	0.352

Notes: (i) Output for GINI (1), Theil (2) and MLD (3). The educational levels medium and high are added and called medhigh. The first value is the coefficient and the value between the brackets is the p-value. (ii) Robust pval in parentheses *** p<0.01, ** p<0.05, * p<0.1.

5.2. Random vs. Fixed Model

First, the fixed effects model and the random effects model are estimated. Then a Hausman test is conducted to see which model should be used. Table 5.6 shows the results of the Hausman test.

	GINI	Theil	MLD
P-value	0.1235	0.1196	0.0091

Table 5.6: P-values for the Hausman test for all three inequality measures

The *GINI* coefficient and the *Theil* index cannot reject the null hypothesis of no systematic differences in coefficients; the p-value exceeds 0.05. Therefore it doesn't matter which model to use. The *MLD* p-value is lower than 0.05, therefore the null hypothesis will be rejected and the Random effects model cannot be used.

Therefore, the fixed effects model is chosen to compare the three inequality measures. The results of this model are shown in Table 5.7. The results of the fixed effects model differ with those of the pooled model. *GDP* is no longer significant at the 5% and 10% level for the *GINI* coefficient and the *Theil* index. It is, however significant at a 5% level for the *MLD*. This is a negative relation, and thus lowers the *MLD*. The Years 2010-2014 are significant for all three inequality measures at a 5% level. All those years have a positive relation with the inequality measures and increase income inequality. *Educ_att_low* and *Educ_att_med* are only significant at a 10% level for all three inequality measures and decreases income inequality. *Educ_att_high* is significant at a 5% level for the *GINI* coefficient and for the *Theil* index. It is only significant at a 10% level for the *MLD*. The relation between *Educ_att_high* and the inequality measures is negative and thus lowers income inequality. The variable *Oldagedependency* is significant at a 5% level and lowers income inequality. The variables *Foreign_country* and *Mortalityrates* are no longer significant at a 5% and 10% level.

Table 5.7: Fixed effects model with three educational levels

VARIABLES	(1) GINI	(2) Theil	(3) MLD
GDP	-2.19e-06	-2.46e-06	-3.76e-06**
	(0.451)	(0.143)	(0.0375)
2007.Year	0.0139	0.00761	0.00973
	(0.202)	(0.228)	(0.150)
2008.Year	0.0185	0.0106	0.0130*
	(0.126)	(0.131)	(0.0822)
2009.Year	0.0210	0.0104	0.0132
	(0.129)	(0.191)	(0.122)
2010.Year	0.0384**	0.0193**	0.0248**
	(0.0187)	(0.0414)	(0.0143)
2011.Year	0.0500***	0.0263**	0.0333***
	(0.00877)	(0.0170)	(0.00484)
2012.Year	0.0619***	0.0331***	0.0405***
	(0.00408)	(0.00786)	(0.00241)
2013.Year	0.0732***	0.0417***	0.0508***
	(0.00466)	(0.00530)	(0.00157)
2014.Year	0.0931***	0.0565***	0.0679***
	(0.00211)	(0.00128)	(0.000311)
Educ_att_low	-9.341*	-5.520*	-5.627*
	(0.0647)	(0.0593)	(0.0722)
Educ_att_med	-9.374*	-5.553*	-5.675*
	(0.0651)	(0.0590)	(0.0712)
Educ_att_high	-9.970**	-5.921**	-6.084*
	(0.0485)	(0.0429)	(0.0517)
Foreign_country	0.0470	0.0241	0.00610
	(0.585)	(0.628)	(0.909)
Oldagedependency	-1.391***	-0.815***	-0.936***
	(0.00167)	(0.00146)	(0.000651)
Mortalityrates	0.785	0.433	0.454
	(0.309)	(0.333)	(0.342)
Constant	9.526*	5.662*	5.835*
	(0.0624)	(0.0556)	(0.0651)
Observations	252	252	252
Number of Country	28	28	28
R-squared	0.135	0.136	0.159

Notes: (i) Output for GINI (1), Theil (2) and MLD (3). The first value is the coefficient and the value between the brackets is the p-value. (ii) Robust pval in parentheses *** p<0.01, ** p<0.05, * p<0.1.

To answer the last partial research question, a spearman rank correlation is conducted. The entire correlation table can be found in Appendix D. The correlation between the three inequality measures is shown in Table 5.8. Any value between 0.8 and 1.0 means a very strong correlation. The correlation between the inequality measures lies between 0.986 and 0.996, this means that the inequality measures are strongly correlated with each other.

	GINI	Theil	MLD
GINI		0.9962	0.9858
Theil	0.9962		0.9861
MLD	0.9858	0.9861	

Table 5.8: Correlation between the inequality measures

6. Conclusion

In this thesis, the relation between income inequality, education and social structure is examined. This relationship is examined for three inequality measures; the *GINI* coefficient, the *Theil* index and the *MLD*.

First, some theory is described on income inequality, education and social structure. Then some earlier research is shown on the subject of income inequality. In further chapters, the data and methodology used are mentioned, and then the results of the models are shown.

The relation is researched in the 28 countries of the EU in the years 2006-2014. Three different models were estimated; the pooled, fixed effects and the random effects model. A Hausman test was conducted to conclude whether to use the fixed effects model or the random effects model. The models are estimated using the least squares methods. When using the robust option in Stata, only the assumption of endogeneity was tested.

From the pooled model we can conclude that a higher level of education does indeed lead to a lower income inequality. The country of birth does also influence income inequality. When the number of foreign-born people working in a country increases, income inequality will increase also. When the aging population increases thus old age dependency increases, income inequality will also increase. Mortality rates are also significant and will decrease income inequality.

According to the fixed effects model, all education levels have a significant effect and decrease income inequality. The country of birth and mortality rates are not significant and have no effect on income inequality. Old age dependency is significant and leads to a decrease in income inequality.

The spearman rank correlation between the three inequality measures is between 0.986 and 0.996, and means that there is a very strong relation between the measures. This is in accordance with the values of Champernowne (1974). In this thesis, there are no significant differences between the income inequality measures. The conclusion that education (Muller, 2002), origin (Barham and Boucher, 1998) and mortality rates (Kawachi et al., 1997) influence income inequality is consistent with prior research.

7. Discussion and Recommendation

The shortcomings will be discussed in this chapter. Also, these shortcomings give room for future research.

The first shortcoming is related to the dataset. The dataset included 28 countries in 9 years with a total of 252 observations. This is not a very big dataset. As mentioned earlier, there are a lot of factors that influence income inequality. In this thesis, only a couple of factors are included, and this can influence the outcome of the research. When there are more factors included, the outcome of this research could be totally different. Also, 9 years are not a lot of years. Unfortunately, data before 2006 is scarce; only a handful of countries had data available.

For further research I would recommend a different measure for the educational levels. When they are split into three categories, they were only significant at a 10% level. But then you can only see that education is significant in the selected model and not if a higher level of education leads to a decrease in income inequality. Splitting the variable *Education* into only two levels: a low and high level will likely resolve this problem.

Horizontal inequality has become more important in recent years. Subjects as racial or social clubs are not included as data is scarce. Another subject that is not included is gender. As only some subjects on horizontal inequality are included, the inequality measurements are complex (Steward, 2000). For future research it is recommended to add more subjects on horizontal inequality.

In the pooled model only the year 2014 is significant to reference year 2006, this is an interesting observation. This means that in 2014, income inequality increases significantly relative to 2006. In the fixed effects model, the years 2010-2014 are significant and income inequality increases significantly. The economic crisis that started in 2007 could be a factor that makes the following years after the crisis significant.

Another interesting observation is that in the pooled model, old age dependency increases income inequality, and in the fixed effects model it decreases income inequality. Whether old age dependency is a factor that increases or decreases income inequality, is a question which researchers do not have the answer yet. In Japan, income inequality has declined while the old age ratio has increased. The population share of the elderly has doubled, while the number of children fell by more than 20% in the last 20 years, according to the OECD. Although recent research has shown that even in Japan, income inequality in elderly households increased.

Also, poverty rates have increased substantially in the last 30 years (Takanami, 2010). But, poverty rates were not included in this thesis. Perhaps, when included, this would mean a different outcome or conclusion.

There are some discrepancies when comparing the pooled model with the fixed model. An explanation for this difference can be that the fixed effects model only allows for omitted variable bias for variables that are constant over time (Brooks, 2014). There are potentially a lot of variables that could have been taken into account, and that these omitted variables are not constant over time. Some examples of other variables are government policy or pension funds. When the government provides social benefits and maintains a minimum wage, this would likely reduce income inequality. This is the same for pension funds. When there are good pension funds, the elderly are less likely to go into poverty, and income inequality will not increase substantially. As mentioned in the introduction the European Union is “united by diversity”, and these diversities must be taken into account.

8. Bibliography

Adams Jr., R. H., and Page, J. (2005). Do International Migration and Remittances Reduce Poverty in Developing Countries? *World Development, 33*(10), pp. 1645-1669.

Antonakis, J., Bendahan, S., Jacquart, P. and Lalive, R. (2014). Causality and Endogeneity: Problems and Solutions. *The Oxford Handbook of Leadership and Organizations.* New York: Oxford University Press, pp. 93-117.

Atkinson, A. (1997). Bringing Income Distribution in From the Cold. *The Economic Journal, 107*, pp. 297-321.

Atkinson, A., and Bourguignon, F. (2000). Intoduction: Income Distribution and Economics. *Handbook of Income Distribution, 1*, pp. 1-58.

Barham, B., and Boucher, S. (1998). Migration, Remittances, and Inequality: Estimating the Net Effects of Migration on Income Distribution. *Journal of Development Economics, 55*(2), pp. 307-331.

Barro, R. J. (2000). Inequality and Growth in a Panel of Countries. *Journal of Economic Growth, 5*, pp. 5-32.

Barro, R. J., and Lee, J. W. (2013). A New Data Set of Educational Attainment in the World, 1950-2010. *Journal of Development Economics, 104*, pp. 184-198.

Becker, G. S. (1964). *Human Capital.* New York: Columbia University Press.

Bentham, G. (1988). Migration and Morbidity: Implications for Geographical Studies of Disease. Journal of Social Science and Medicine, 26(1), pp. 49-54.

Blaug, M. (1976). The Empirical Status of Human Capital Theory: A Slightly Jaundiced Survey. *Journal of Economic Literature, 14*(3), pp. 827-855.

Breen, R., and Chung, I. (2015). Income Inequality and Education. *Sociological Science, 2*, pp. 454-477.

Brooks, C. (2014). *Introductory Econometrics for Finance* (3rd Edition ed.). Cambridge University Press.

Brown, R. L., and Prus, S. G. (2004). Social Transfers and Income Inequality in Old-age: A Multi-national Perspective? *North American Actuarial Journal, 8*(4), pp. 30-36.

Carter Hill, R., Griffiths, W. E., and Lim, G. C. (2012). *Principles of Econometrics* (4 ed.). Hoboken: Wiley.

Champernowne, D. G. (1974). A Comparison of Measures of Inequality of Income Distribution. *The Economic Journal, 84*(336), pp. 787-816.

Chiswick, B. R. (1971). Earnings Inequality and Economic Development. *The Quarterly Journal of Economics, 85*(1), pp. 21-39.

Conceição, P., and Ferreira, P. (2000). The Young Person's Guide to the Theil Index: Suggesting Intuitive Interpretations and Exploring Analytical Applications. *Working Paper*.

De Gregorio, J., and Lee, J.-W. (2002). Education and Income Inequality: New Evidence From Cross-Country Data. *Review of Income and Wealth, 48*(3), 395-416.

Di Falco, E. (2014). *Income Inequality Statistics.* Retrieved February 29, 2016, from Eurostat: http://ec.europa.eu/eurostat/statistics-explained/index.php/Income_inequality_statistics

Eurostat. (2016). *Gini Coefficient of Equavalised Disposable Income.* Retrieved February 26, 2016, from Eurostat: http://appsso.eurostat.ec.europa.eu/nui/submitViewTableAction.do

Fredriksen, K. B. (2012). Income Inequality in the European Union. OECD, *Economic Department Working Papers No. 952.*

Gastwirth, J. L. (1972). The Estimation of the Lorenz Curve and Gini Index. *The Review of Economics and Statistics, 54*(3), pp. 306-316.

Gupta, S., Davoodi, H., and Alonso-Terme, R. (2002). Does Corruption Affect Income Inequality and Poverty? *Economics of Governance, 3*(1), pp. 23-45.

Hausman, J. A. (1978). Specification Test in Econometrics. *Econometrica 46,* pp. 1251-1271.

Hoover, E. M. (1936). The Measurement of Industrial Localization. *The Review of Economics and Statistics, 18*(4), pp. 162-171.

Kaplan, G., Pamuk, E., Lynch, J., Cohen, R., and Balfour, J. (1996). Inequality in Income and Mortality in the United States: Analysis of Mortality and Potential Pathways. *BMJ, 312.*

Kawachi, I., and Kennedy, B. P. (1997). The Relationship of Income Inequality to Mortality: Does the Choice of Indicator Matter? *Social Science and Medicine, 45*(7), pp. 1121-1127.

Kawachi, I., Kennedy, B. P., Lochner, K., and Prothrow-Stith, D. (1997). Social Capital, Income Inequality, and Mortality. *American Journal of Public Health, 87*(9), pp. 1491-1498.

Kennedy, B. P., Kawachi, I., and Prothrow-Stith, D. (1996). Income distribution and Mortality: Cross-sectional Ecological Study of the Robin Hood Index in the United States. *BMJ,312*, pp. 1004-1007

Kennedy, B. P., Kawachi, I., Prothrow-Stith, D., Lochner, K., and Gupta, V. (1998). Social Capital, Income Inequality, and Firearm Violent Crime. *Social Science and Medicine, 47*(1), pp. 7-17.

Krugman, P. (2016, June 11). Hillary and the Horizontals. *International New York Times*, pp. 7.

Kuznets, S. (1955). Economic Growth and Income Inequality. *The American Economic Review, 45*(1), pp. 1-28.

Liebig, T., and Sousa-Poza, A. (2004). Migration, Self-Selection and Income Inequality: An International Analysis. *Kyklos: International Review for Social Sciences, 57*(1), pp. 125-146.

Lleras-Muney, A. (2005). The Relationship Between Education and Adult Mortality in the United States. *Review of Economic Studies, 72*, pp. 189-221.

Lorgelly, P. K., and Lindley, J. (2008). What is the Relationship Between Income Inequality and Health? Evidence from the BHPS. *Health Economics, 17*, pp. 249-265.

McKenzie, D., and Rapoport, H. (2006). Can Migration Reduce Educational Attainment? *World Bank Policy Research Working Paper 3952.*

Mendelson, A. (2011). *Inequality and Health.* Retrieved June 12, 2016, from Inequality.org: http://inequality.org/inequality-health/

Muller, A. (2002). Education, Income Inequality, And Mortality: A Multiple Regression Analysis. *British Medical Journal, 324*(7328), pp. 23-25.

Mundalk, Y. (1978). On the Pooling of Time Series and Cross-section Data. *Econometrica, 46*, pp. 69-85.

Nelson, R. R., and Phelps, E. S. (1966). Investment in Humans, Technological Diffusion, and Economic growth. *The American Economic Review, 56*(1/2), pp. 69-75.

OECD. (2015). *Government at a glance 2015.* Paris: OECD Publishing.

OECD. *Inequality and income*. Retrieved January 29, 2016, from OECD: http://www.oecd.org/social/inequality.htm

Oshio, T., and Kobayashi, M. (2010). Income Inequality, Perceived Happiness, and Self-rated Health: Evidence From Nationwide Surveys in Japan. *Social Science and Medicine, 70*(9), pp. 1358-1366.

Pettigrew, T. F., Coleman, J. S., Sewell, W. H., and Pullum, T. W. (1973). Review of Inequality: A Reassessment of the Effect of Family and Schooling in America. *Journal of Sociology, 78*(6), pp. 1523-1544.

Ray, D. (1998). *Development Economics.* Princeton University Press.

Radcliffe-Brown, A.R. (1940). On Social Structure. *The Journal of the Royal Anthropological Institute of Great Britain and Ireland, 70* (1), pp. 1-12.

Rodgers, G. (1979). Income and Inequality as Determinants of Mortality: An International Cross-Section Analysis. *Popultaion Studies, 33*(2), pp. 343-351.

Sala-i-Martin, X. (2002). The World Distribution of Income (Estimated From Individual Country Distribution), *NBER Working Paper* 8933.

Sarel, M. (1997). How Macroeconomic Factors Affect Income Distribution: The Cross-Country Evidence. *IMF Working Paper*, International Monetary Fund, Asia and Pacific Department.

Schultz, T. W. (1961). Investment in Human Capital. *American Economic Review, 51*, pp.1-17.

Stark, O., Taylor, J. E., and Yitzhaki, S. (1988). Migration, Remittances and Inequality: A Sensitivity Analysis Using the Extended Gini Index. *Journal of Development Economics, 28* (3), pp. 309-322.

Steward, F. (2000). Crisis Prevention: Tackling Horizontal Inequalities. *Oxford Development Studies, 28* (3), pp. 245-262.

Steward, F., Brown, G., and Mancini, L. (2005). Why Horizontal Inequalities Matter: Some Implications for Measurement. Crise Working Paper No. 19.

Takanami, K. (2010). Rising income inequality and poverty in Japan. Master thesis, San Diego State University, Liberal Arts and Sciences.

Thomas, V., Wang, Y., and Fan, X. (2001): Measuring Education Inequality: Gini Coefficients of Education. World Bank Publications.

Walker, A. (1980). The Social Creation of Poverty and Dependency in Old Age. *Journal of Social Policy, 9* (1), pp. 49-75.

Weich, S., Lewis, G., and Jenkins, S. P. (2002). Income Inequality and Self-rated Health in Britain. *Epidemiol Community Health, 56* (6), pp. 436-441.

Wilkinson, R. G., and Pickett, K. E. (2006). Income Inequality and Population Health: A Review and Explanation of the Evidence. *Social Science and Medicine, 62*(7), pp. 1768-1784.

Williams, R. (2015). Heteroskedasticity. Retrieved May 25, 2016, from University of Notre Dame: https://www3.nd.edu/~rwilliam/stats2/l25.pdf

Willis, R. J., and Rosen, S. (1978). Education and Self-Selection, *NBER Working Paper 249.*

9. Appendix

Appendix A: List of EU countries

Country	Country Number
Austria	1
Belgium	2
Bulgaria	3
Croatia	4
Cyprus	5
Czech Republic	6
Denmark	7
Estonia	8
Finland	9
France	10
Germany	11
Greece	12
Hungary	13
Ireland	14
Italy	15
Latvia	16
Lithuania	17
Luxembourg	18
Malta	19
Netherlands	20
Poland	21
Portugal	22
Romania	23
Slovakia	24
Slovenia	25
Spain	26
Sweden	27
United Kingdom	28

Appendix B: Data used for EU

Country	Year	GINI	Theil	MLD	GDP	Educ_att_low	Educ_att_ med	Educ_att_ med+high	Educ_att_high	Foreign_country	Old-age dependency	Mortality rates
1	2006	0,253	0,098	0,104	€ 34.500	0,250	0,603	0,750	0,147	0,147	0,243	0,801
1	2007	0,262	0,109	0,107	€ 35.700	0,254	0,600	0,746	0,146	0,150	0,250	0,803
1	2008	0,277	0,121	0,121	€ 36.100	0,246	0,604	0,754	0,150	0,153	0,254	0,806
1	2009	0,275	0,121	0,123	€ 34.700	0,238	0,602	0,762	0,160	0,155	0,258	0,805
1	2010	0,283	0,127	0,130	€ 35.200	0,232	0,605	0,767	0,162	0,150	0,262	0,807
1	2011	0,274	0,118	0,122	€ 36.100	0,231	0,606	0,769	0,163	0,161	0,261	0,811
1	2012	0,276	0,120	0,125	€ 36.200	0,224	0,608	0,777	0,169	0,162	0,263	0,811
1	2013	0,270	0,115	0,121	€ 36.100	0,222	0,602	0,779	0,177	0,167	0,268	0,813
1	2014	0,276	0,120	0,123	€ 36.000	0,203	0,524	0,798	0,274	0,167	0,272	0,817
2	2006	0,278	0,120	0,126	€ 33.100	0,357	0,364	0,643	0,279	0,125	0,262	0,795
2	2007	0,263	0,108	0,110	€ 34.000	0,348	0,371	0,652	0,281	0,130	0,259	0,799
2	2008	0,275	0,118	0,121	€ 34.000	0,336	0,380	0,664	0,284	0,135	0,258	0,798
2	2009	0,264	0,108	0,110	€ 32.900	0,324	0,381	0,675	0,294	0,139	0,259	0,801
2	2010	0,266	0,108	0,114	€ 33.500	0,326	0,367	0,674	0,307	0,149	0,260	0,803
2	2011	0,263	0,107	0,111	€ 33.900	0,319	0,377	0,681	0,304	0,150	0,260	0,807
2	2012	0,265	0,109	0,114	€ 33.700	0,314	0,374	0,687	0,313	0,153	0,264	0,805
2	2013	0,259	0,104	0,106	€ 33.500	0,304	0,381	0,696	0,315	0,155	0,268	0,807
2	2014	0,259	0,102	0,107	€ 33.800	0,295	0,378	0,704	0,326	0,158	0,273	0,814
3	2006	0,312	0,153	0,161	€ 4.600	0,306	0,513	0,695	0,182	0,000	0,253	0,727
3	2007	0,353	0,198	0,222	€ 4.900	0,287	0,528	0,713	0,185	0,000	0,255	0,73
3	2008	0,359	0,204	0,212	€ 5.300	0,283	0,528	0,717	0,189	0,000	0,258	0,733
3	2009	0,334	0,175	0,184	€ 5.100	0,276	0,532	0,724	0,192	0,000	0,261	0,737
3	2010	0,332	0,173	0,182	€ 5.100	0,259	0,547	0,741	0,194	0,000	0,265	0,738
3	2011	0,350	0,191	0,207	€ 5.200	0,240	0,559	0,760	0,201	0,011	0,270	0,742
3	2012	0,336	0,178	0,191	€ 5.300	0,230	0,564	0,771	0,207	0,012	0,278	0,744
3	2013	0,354	0,200	0,211	€ 5.400	0,221	0,557	0,779	0,222	0,013	0,285	0,749

3	2014	0,354	0,201	0,214	€ 5.500	0,226	0,538	0,774	0,236	0,015	0,293	0,745
4	2006	0,000	0,000	0,000	€ 10.700	0,283	0,582	0,718	0,136	0,000	0,263	0,759
4	2007	0,000	0,000	0,000	€ 11.200	0,267	0,599	0,733	0,134	0,000	0,265	0,758
4	2008	0,000	0,000	0,000	€ 11.500	0,259	0,605	0,741	0,136	0,000	0,267	0,76
4	2009	0,000	0,000	0,000	€ 10.600	0,253	0,602	0,747	0,145	0,000	0,268	0,763
4	2010	0,316	0,157	0,171	€ 10.500	0,254	0,588	0,745	0,157	0,000	0,267	0,767
4	2011	0,312	0,152	0,167	€ 10.500	0,246	0,600	0,754	0,154	0,000	0,265	0,772
4	2012	0,309	0,148	0,161	€ 10.300	0,231	0,611	0,769	0,158	0,000	0,267	0,773
4	2013	0,309	0,150	0,161	€ 10.200	0,224	0,605	0,775	0,170	0,128	0,271	0,778
4	2014	0,302	0,142	0,153	€ 10.200	0,208	0,607	0,792	0,185	0,134	0,276	0,779
5	2006	0,288	0,131	0,130	€ 23.600	0,336	0,388	0,664	0,276	0,000	0,178	0,801
5	2007	0,298	0,143	0,137	€ 24.200	0,311	0,391	0,688	0,297	0,000	0,180	0,798
5	2008	0,290	0,131	0,132	€ 24.500	0,304	0,386	0,696	0,310	0,000	0,179	0,806
5	2009	0,295	0,140	0,133	€ 23.300	0,312	0,382	0,687	0,305	0,189	0,178	0,81
5	2010	0,301	0,142	0,142	€ 23.000	0,296	0,383	0,704	0,321	0,230	0,178	0,815
5	2011	0,292	0,133	0,134	€ 22.600	0,283	0,380	0,717	0,337	0,231	0,180	0,812
5	2012	0,310	0,154	0,148	€ 21.700	0,264	0,386	0,736	0,350	0,232	0,181	0,811
5	2013	0,324	0,168	0,163	€ 20.400	0,254	0,392	0,746	0,354	0,232	0,188	0,825
5	2014	0,348	0,199	0,186	€ 20.100	0,261	0,375	0,739	0,364	0,223	0,199	0,828
6	2006	0,253	0,102	0,099	€ 14.400	0,164	0,722	0,836	0,114	0,055	0,200	0,767
6	2007	0,253	0,102	0,097	€ 15.200	0,162	0,722	0,838	0,116	0,062	0,203	0,77
6	2008	0,247	0,097	0,095	€ 15.400	0,158	0,717	0,841	0,124	0,065	0,206	0,773
6	2009	0,251	0,103	0,098	€ 14.600	0,152	0,714	0,848	0,134	0,064	0,211	0,774
6	2010	0,249	0,100	0,097	€ 14.900	0,144	0,711	0,856	0,145	0,063	0,217	0,777
6	2011	0,252	0,103	0,097	€ 15.200	0,139	0,703	0,861	0,158	0,071	0,223	0,78
6	2012	0,249	0,100	0,096	€ 15.000	0,134	0,696	0,866	0,170	0,071	0,234	0,781
6	2013	0,246	0,094	0,097	€ 15.000	0,129	0,690	0,871	0,181	0,071	0,246	0,783
6	2014	0,251	0,102	0,097	€ 15.200	0,124	0,685	0,876	0,191	0,073	0,257	0,789
7	2006	0,237	0,088	0,095	€ 46.000	0,254	0,453	0,746	0,293	0,066	0,229	0,784
7	2007	0,252	0,103	0,107	€ 46.200	0,323	0,417	0,677	0,260	0,069	0,232	0,784

7	2008	0,251	0,100	0,100	€ 45.600	0,329	0,408	0,671	0,263	0,073	0,236	0,788
7	2009	0,269	0,125	0,170	€ 43.000	0,322	0,409	0,678	0,269	0,075	0,241	0,79
7	2010	0,269	0,120	0,150	€ 43.500	0,318	0,407	0,682	0,275	0,077	0,249	0,793
7	2011	0,278	0,113	0,117	€ 43.900	0,307	0,415	0,694	0,279	0,079	0,257	0,799
7	2012	0,281	0,111	0,116	€ 43.700	0,298	0,416	0,702	0,286	0,082	0,267	0,802
7	2013	0,268	0,116	0,117	€ 43.400	0,295	0,414	0,705	0,291	0,085	0,276	0,804
7	2014	0,277	0,124	0,123	€ 43.700	0,282	0,420	0,718	0,298	0,087	0,283	0,807
8	2006	0,331	0,173	0,178	€ 12.200	0,206	0,520	0,794	0,274	0,168	0,248	0,732
8	2007	0,334	0,176	0,178	€ 13.300	0,200	0,525	0,800	0,275	0,167	0,255	0,732
8	2008	0,309	0,147	0,158	€ 12.600	0,204	0,513	0,796	0,283	0,166	0,258	0,744
8	2009	0,314	0,153	0,160	€ 10.800	0,192	0,506	0,808	0,302	0,163	0,258	0,753
8	2010	0,313	0,151	0,159	€ 11.000	0,182	0,518	0,818	0,300	0,160	0,259	0,76
8	2011	0,319	0,159	0,169	€ 11.900	0,177	0,510	0,823	0,313	0,159	0,260	0,766
8	2012	0,325	0,165	0,175	€ 12.600	0,168	0,511	0,832	0,321	0,100	0,265	0,767
8	2013	0,329	0,171	0,177	€ 12.800	0,158	0,519	0,842	0,323	0,101	0,272	0,775
8	2014	0,356	0,199	0,211	€ 13.200	0,151	0,522	0,848	0,326	0,102	0,279	0,774
9	2006	0,259	0,107	0,103	€ 35.500	0,262	0,448	0,738	0,290	0,036	0,240	0,795
9	2007	0,262	0,109	0,104	€ 37.200	0,254	0,446	0,746	0,300	0,038	0,248	0,796
9	2008	0,263	0,110	0,107	€ 37.300	0,251	0,446	0,748	0,302	0,041	0,248	0,799
9	2009	0,259	0,105	0,104	€ 34.000	0,244	0,447	0,756	0,309	0,044	0,252	0,801
9	2010	0,254	0,102	0,099	€ 34.900	0,236	0,448	0,764	0,316	0,046	0,256	0,802
9	2011	0,258	0,103	0,105	€ 35.600	0,229	0,446	0,771	0,325	0,049	0,265	0,806
9	2012	0,259	0,103	0,105	€ 34.900	0,218	0,454	0,782	0,328	0,053	0,277	0,807
9	2013	0,254	0,100	0,102	€ 34.500	0,208	0,456	0,792	0,336	0,056	0,289	0,811
9	2014	0,256	0,102	0,100	€ 34.100	0,201	0,452	0,799	0,347	0,055	0,302	0,813
10	2006	0,273	0,117	0,117	€ 31.000	0,350	0,411	0,651	0,240	0,114	0,251	0,809
10	2007	0,266	0,110	0,111	€ 31.500	0,340	0,416	0,660	0,244	0,115	0,251	0,813
10	2008	0,298	0,146	0,138	€ 31.400	0,330	0,423	0,671	0,248	0,116	0,252	0,814
10	2009	0,299	0,146	0,139	€ 30.300	0,324	0,418	0,677	0,259	0,116	0,254	0,815
10	2010	0,298	0,143	0,139	€ 30.800	0,319	0,419	0,681	0,262	0,117	0,256	0,818

10	2011	0,308	0,154	0,149	€ 31.200	0,312	0,420	0,687	0,267	0,119	0,259	0,823
10	2012	0,305	0,152	0,141	€ 31.200	0,304	0,419	0,696	0,277	0,119	0,267	0,821
10	2013	0,301	0,146	0,143	€ 31.200	0,280	0,431	0,720	0,289	0,115	0,275	0,824
10	2014	0,292	0,136	0,134	€ 31.100	0,265	0,437	0,734	0,297	0,116	0,284	0,828
11	2006	0,268	0,115	0,122	€ 31.000	0,246	0,553	0,754	0,201	0,127	0,289	0,799
11	2007	0,304	0,148	0,156	€ 32.100	0,235	0,561	0,765	0,204	0,128	0,299	0,801
11	2008	0,302	0,145	0,149	€ 32.500	0,223	0,563	0,777	0,214	0,129	0,304	0,802
11	2009	0,291	0,134	0,134	€ 30.800	0,220	0,558	0,781	0,223	0,129	0,309	0,803
11	2010	0,293	0,134	0,135	€ 32.100	0,214	0,561	0,787	0,226	0,130	0,314	0,805
11	2011	0,290	0,131	0,135	€ 33.300	0,181	0,576	0,819	0,243	0,120	0,312	0,808
11	2012	0,283	0,125	0,126	€ 33.400	0,179	0,572	0,821	0,249	0,124	0,312	0,81
11	2013	0,297	0,139	0,141	€ 33.400	0,177	0,571	0,823	0,252	0,128	0,313	0,809
11	2014	0,307	0,150	0,163	€ 33.800	0,197	0,571	0,803	0,232	0,122	0,315	0,812
12	2006	0,343	0,187	0,194	€ 22.000	0,414	0,399	0,586	0,187	0,000	0,277	0,798
12	2007	0,343	0,187	0,196	€ 22.700	0,408	0,401	0,592	0,191	0,000	0,279	0,797
12	2008	0,334	0,176	0,187	€ 22.600	0,397	0,405	0,603	0,198	0,000	0,280	0,802
12	2009	0,331	0,176	0,181	€ 21.500	0,396	0,405	0,604	0,199	0,118	0,282	0,804
12	2010	0,329	0,171	0,180	€ 20.300	0,385	0,406	0,615	0,209	0,119	0,286	0,806
12	2011	0,335	0,175	0,195	€ 18.500	0,371	0,407	0,629	0,222	0,119	0,292	0,808
12	2012	0,343	0,187	0,212	€ 17.200	0,359	0,413	0,642	0,229	0,118	0,300	0,807
12	2013	0,344	0,187	0,207	€ 16.800	0,345	0,414	0,654	0,240	0,116	0,307	0,814
12	2014	0,345	0,186	0,203	€ 17.000	0,331	0,422	0,668	0,246	0,116	0,316	0,815
13	2006	0,333	0,182	0,181	€ 10.200	0,273	0,577	0,727	0,150	0,034	0,229	0,735
13	2007	0,256	0,102	0,104	€ 10.300	0,263	0,583	0,737	0,154	0,038	0,232	0,736
13	2008	0,252	0,099	0,099	€ 10.400	0,258	0,578	0,742	0,164	0,039	0,235	0,742
13	2009	0,247	0,096	0,096	€ 9.700	0,250	0,581	0,750	0,169	0,041	0,238	0,744
13	2010	0,241	0,088	0,093	€ 9.800	0,244	0,585	0,756	0,171	0,045	0,242	0,747
13	2011	0,269	0,113	0,114	€ 10.000	0,243	0,577	0,757	0,180	0,047	0,244	0,751
13	2012	0,272	0,116	0,116	€ 9.900	0,235	0,575	0,765	0,190	0,043	0,246	0,753
13	2013	0,283	0,124	0,128	€ 10.100	0,228	0,578	0,773	0,195	0,045	0,251	0,758

13	2014	0,286	0,127	0,133	€ 10.500	0,220	0,577	0,779	0,202	0,045	0,258	0,76
14	2006	0,319	0,161	0,155	€ 39.800	0,352	0,370	0,647	0,277	0,142	0,160	0,793
14	2007	0,313	0,153	0,151	€ 40.700	0,341	0,370	0,659	0,289	0,156	0,157	0,797
14	2008	0,299	0,140	0,138	€ 39.000	0,324	0,373	0,676	0,303	0,165	0,156	0,802
14	2009	0,288	0,127	0,132	€ 36.500	0,315	0,371	0,685	0,314	0,169	0,160	0,802
14	2010	0,307	0,149	0,149	€ 36.400	0,305	0,367	0,694	0,327	0,170	0,165	0,808
14	2011	0,298	0,140	0,143	€ 37.200	0,297	0,370	0,703	0,333	0,164	0,172	0,809
14	2012	0,299	0,140	0,146	€ 37.200	0,288	0,365	0,712	0,347	0,163	0,179	0,809
14	2013	0,300	0,142	0,142	€ 37.600	0,272	0,364	0,727	0,363	0,164	0,186	0,811
14	2014	0,308	0,151	0,150	€ 39.500	0,256	0,386	0,744	0,358	0,161	0,193	0,814
15	2006	0,321	0,162	0,174	€ 28.500	0,495	0,391	0,505	0,114	0,000	0,301	0,814
15	2007	0,320	0,162	0,169	€ 28.700	0,487	0,393	0,513	0,120	0,000	0,305	0,816
15	2008	0,312	0,152	0,164	€ 28.200	0,479	0,395	0,522	0,127	0,074	0,307	0,817
15	2009	0,318	0,158	0,169	€ 26.500	0,471	0,401	0,529	0,128	0,099	0,309	0,818
15	2010	0,317	0,158	0,167	€ 26.800	0,463	0,407	0,537	0,130	0,098	0,312	0,822
15	2011	0,325	0,166	0,183	€ 26.900	0,455	0,414	0,546	0,132	0,097	0,313	0,824
15	2012	0,324	0,166	0,179	€ 26.000	0,442	0,419	0,558	0,139	0,096	0,320	0,824
15	2013	0,328	0,171	0,188	€ 25.400	0,434	0,422	0,566	0,144	0,095	0,327	0,829
15	2014	0,324	0,167	0,185	€ 25.300	0,423	0,427	0,577	0,150	0,094	0,331	0,832
16	2006	0,389	0,243	0,252	€ 9.200	0,246	0,580	0,754	0,174	0,000	0,248	0,706
16	2007	0,354	0,196	0,206	€ 10.200	0,240	0,575	0,760	0,185	0,000	0,254	0,708
16	2008	0,375	0,222	0,237	€ 9.900	0,228	0,565	0,772	0,207	0,153	0,257	0,721
16	2009	0,375	0,221	0,239	€ 8.600	0,215	0,570	0,784	0,214	0,150	0,262	0,728
16	2010	0,359	0,203	0,224	€ 8.500	0,195	0,578	0,804	0,226	0,148	0,268	0,731
16	2011	0,351	0,194	0,211	€ 9.200	0,195	0,568	0,804	0,236	0,146	0,272	0,739
16	2012	0,357	0,204	0,210	€ 9.700	0,174	0,574	0,826	0,252	0,141	0,276	0,741
16	2013	0,352	0,193	0,209	€ 10.000	0,166	0,564	0,834	0,270	0,138	0,281	0,743
16	2014	0,355	0,201	0,212	€ 10.400	0,161	0,570	0,839	0,269	0,135	0,288	0,745
17	2006	0,350	0,195	0,205	€ 8.700	0,207	0,574	0,794	0,220	0,058	0,243	0,71
17	2007	0,338	0,181	0,191	€ 9.800	0,199	0,564	0,801	0,237	0,057	0,247	0,707

17	2008	0,345	0,188	0,197	€ 10.100	0,185	0,561	0,814	0,253	0,055	0,252	0,717
17	2009	0,359	0,205	0,208	€ 8.700	0,180	0,565	0,820	0,255	0,054	0,254	0,729
17	2010	0,370	0,219	0,234	€ 9.000	0,171	0,560	0,829	0,269	0,051	0,256	0,733
17	2011	0,330	0,172	0,187	€ 9.800	0,159	0,563	0,842	0,279	0,049	0,266	0,737
17	2012	0,320	0,160	0,169	€ 10.300	0,150	0,564	0,850	0,286	0,048	0,269	0,741
17	2013	0,346	0,189	0,195	€ 10.800	0,144	0,558	0,856	0,298	0,047	0,272	0,741
17	2014	0,350	0,194	0,199	€ 11.200	0,142	0,543	0,857	0,314	0,047	0,275	0,747
18	2006	0,278	0,124	0,127	€ 77.800	0,393	0,402	0,607	0,205	0,371	0,208	0,794
18	2007	0,274	0,116	0,119	€ 82.900	0,387	0,386	0,613	0,227	0,373	0,207	0,795
18	2008	0,277	0,122	0,123	€ 80.800	0,368	0,395	0,632	0,237	0,398	0,206	0,807
18	2009	0,292	0,135	0,131	€ 75.100	0,285	0,413	0,715	0,302	0,396	0,205	0,808
18	2010	0,279	0,120	0,125	€ 77.900	0,289	0,408	0,711	0,303	0,405	0,204	0,808
18	2011	0,272	0,115	0,115	€ 78.100	0,291	0,393	0,710	0,317	0,415	0,203	0,811
18	2012	0,280	0,122	0,124	€ 75.600	0,284	0,382	0,716	0,334	0,426	0,203	0,815
18	2013	0,304	0,145	0,144	€ 76.900	0,261	0,386	0,738	0,352	0,437	0,202	0,819
18	2014	0,287	0,129	0,135	€ 78.200	0,249	0,355	0,751	0,396	0,423	0,204	0,823
19	2006	0,271	0,115	0,116	€ 15.000	0,675	0,215	0,326	0,111	0,000	0,199	0,795
19	2007	0,263	0,108	0,112	€ 15.500	0,673	0,210	0,327	0,117	0,000	0,199	0,799
19	2008	0,281	0,122	0,129	€ 16.000	0,651	0,228	0,349	0,121	0,000	0,199	0,797
19	2009	0,274	0,117	0,116	€ 15.500	0,620	0,252	0,380	0,128	0,074	0,203	0,804
19	2010	0,286	0,127	0,130	€ 15.900	0,603	0,256	0,398	0,142	0,080	0,214	0,815
19	2011	0,272	0,115	0,116	€ 16.200	0,589	0,260	0,411	0,151	0,081	0,227	0,809
19	2012	0,271	0,114	0,114	€ 16.500	0,567	0,269	0,433	0,164	0,084	0,239	0,809
19	2013	0,279	0,119	0,124	€ 17.000	0,549	0,279	0,451	0,172	0,089	0,251	0,819
19	2014	0,277	0,120	0,120	€ 17.500	0,536	0,284	0,464	0,180	0,094	0,264	0,821
20	2006	0,264	0,109	0,113	€ 37.600	0,323	0,415	0,677	0,262	0,106	0,211	0,8
20	2007	0,276	0,121	0,121	€ 38.900	0,315	0,419	0,686	0,267	0,107	0,215	0,804
20	2008	0,276	0,123	0,120	€ 39.400	0,313	0,408	0,686	0,278	0,109	0,218	0,805
20	2009	0,272	0,118	0,117	€ 37.700	0,312	0,404	0,688	0,284	0,111	0,223	0,809
20	2010	0,255	0,102	0,103	€ 38.000	0,319	0,404	0,681	0,277	0,112	0,228	0,81

20	2011	0,258	0,104	0,109	€ 38.500	0,316	0,404	0,684	0,280	0,114	0, 233	0,813
20	2012	0,254	0,100	0,103	€ 37.900	0,308	0,406	0,692	0,286	0,115	0,244	0,812
20	2013	0,251	0,100	0,099	€ 37.600	0,290	0,416	0,709	0,293	0,116	0,255	0,814
20	2014	0,262	0,108	0,111	€ 37.900	0,288	0,414	0,711	0,297	0,116	0,264	0,818
21	2006	0,333	0,175	0,181	€ 8.000	0,210	0,642	0,791	0,149	0,000	0,189	0,753
21	2007	0,322	0,163	0,168	€ 8.600	0,204	0,639	0,796	0,157	0,000	0,190	0,754
21	2008	0,320	0,163	0,164	€ 8.900	0,196	0,638	0,803	0,165	0,000	0,189	0,756
21	2009	0,314	0,155	0,158	€ 9.100	0,187	0,632	0,813	0,181	0,012	0,189	0,759
21	2010	0,311	0,153	0,153	€ 9.400	0,180	0,626	0,820	0,194	0,017	0,191	0,764
21	2011	0,311	0,154	0,155	€ 9.900	0,175	0,622	0,825	0,203	0,017	0,191	0,768
21	2012	0,309	0,150	0,152	€ 10.000	0,167	0,618	0,833	0,215	0,017	0,197	0,769
21	2013	0,307	0,149	0,148	€ 10.100	0,161	0,613	0,839	0,226	0,016	0,204	0,771
21	2014	0,308	0,149	0,153	€ 10.500	0,155	0,607	0,845	0,238	0,016	0,212	0,778
22	2006	0,377	0,232	0,231	€ 16.800	0,717	0,167	0,284	0,117	0,072	0,260	0,79
22	2007	0,368	0,218	0,215	€ 17.200	0,714	0,166	0,286	0,120	0,073	0,263	0,793
22	2008	0,358	0,205	0,204	€ 17.200	0,708	0,166	0,292	0,126	0,075	0,266	0,795
22	2009	0,354	0,202	0,199	€ 16.700	0,692	0,177	0,308	0,131	0,079	0,270	0,797
22	2010	0,337	0,183	0,180	€ 17.000	0,673	0,188	0,327	0,139	0,081	0,275	0,801
22	2011	0,342	0,190	0,186	€ 16.700	0,642	0,203	0,358	0,155	0,083	0,282	0,807
22	2012	0,345	0,190	0,192	€ 16.100	0,616	0,217	0,384	0,167	0,081	0,288	0,806
22	2013	0,342	0,187	0,189	€ 16.000	0,594	0,230	0,406	0,176	0,084	0,294	0,809
22	2014	0,345	0,190	0,198	€ 16.300	0,563	0,240	0,437	0,197	0,082	0,303	0,813
23	2006	0,000	0,000	0,000	€ 5.600	0,321	0,584	0,680	0,096	0,000	0,216	0,725
23	2007	0,378	0,228	0,249	€ 6.100	0,309	0,591	0,690	0,099	0,000	0,215	0,731
23	2008	0,360	0,203	0,222	€ 6.700	0,302	0,592	0,699	0,107	0,000	0,226	0,735
23	2009	0,349	0,191	0,208	€ 6.300	0,302	0,586	0,698	0,112	0,008	0,237	0,737
23	2010	0,333	0,173	0,185	€ 6.300	0,306	0,575	0,694	0,119	0,000	0,237	0,737
23	2011	0,332	0,172	0,193	€ 6.400	0,300	0,572	0,701	0,129	0,000	0,237	0,744
23	2012	0,332	0,174	0,193	€ 6.400	0,292	0,573	0,708	0,135	0,009	0,237	0,744
23	2013	0,340	0,180	0,208	€ 6.700	0,289	0,573	0,711	0,138	0,009	0,239	0,752

23	2014	0,347	0,188	0,221	€ 6.900	0,312	0,546	0,688	0,142	0,011	0,243	0,75
24	2006	0,281	0,128	0,125	€ 10.700	0,189	0,692	0,811	0,119	0,056	0,165	0,745
24	2007	0,245	0,095	0,094	€ 11.900	0,184	0,697	0,816	0,119	0,068	0,167	0,746
24	2008	0,237	0,087	0,089	€ 12.600	0,176	0,701	0,824	0,123	0,082	0,168	0,749
24	2009	0,248	0,099	0,099	€ 11.800	0,165	0,701	0,835	0,134	0,029	0,170	0,753
24	2010	0,259	0,106	0,108	€ 12.400	0,163	0,687	0,838	0,151	0,032	0,173	0,756
24	2011	0,257	0,104	0,108	€ 12.800	0,157	0,680	0,844	0,164	0,026	0,175	0,761
24	2012	0,253	0,099	0,104	€ 13.000	0,150	0,680	0,850	0,170	0,027	0,178	0,763
24	2013	0,242	0,092	0,098	€ 13.200	0,147	0,675	0,852	0,177	0,029	0,184	0,766
24	2014	0,261	0,108	0,113	€ 13.500	0,152	0,667	0,848	0,181	0,029	0,190	0,77
25	2006	0,237	0,089	0,089	€ 17.500	0,228	0,594	0,772	0,178	0,000	0,222	0,783
25	2007	0,232	0,083	0,087	€ 18.600	0,222	0,593	0,778	0,185	0,000	0,227	0,784
25	2008	0,234	0,085	0,088	€ 19.200	0,219	0,592	0,782	0,190	0,000	0,233	0,791
25	2009	0,227	0,081	0,083	€ 17.500	0,208	0,596	0,792	0,196	0,000	0,236	0,794
25	2010	0,238	0,087	0,090	€ 17.700	0,209	0,589	0,791	0,202	0,112	0,238	0,798
25	2011	0,238	0,090	0,091	€ 17.800	0,197	0,588	0,804	0,216	0,112	0,239	0,801
25	2012	0,237	0,088	0,092	€ 17.300	0,190	0,580	0,810	0,230	0,146	0,244	0,803
25	2013	0,244	0,094	0,096	€ 17.100	0,185	0,571	0,815	0,244	0,161	0,250	0,805
25	2014	0,250	0,098	0,100	€ 17.600	0,184	0,565	0,816	0,251	0,161	0,257	0,812
26	2006	0,319	0,160	0,175	€ 24.100	0,509	0,228	0,492	0,264	0,118	0,242	0,811
26	2007	0,319	0,160	0,174	€ 24.500	0,501	0,231	0,499	0,268	0,134	0,240	0,811
26	2008	0,324	0,164	0,176	€ 24.400	0,498	0,231	0,501	0,270	0,141	0,238	0,815
26	2009	0,329	0,172	0,187	€ 23.300	0,495	0,232	0,506	0,274	0,142	0,241	0,819
26	2010	0,335	0,176	0,196	€ 23.200	0,481	0,234	0,518	0,284	0,143	0,246	0,824
26	2011	0,340	0,185	0,203	€ 22.900	0,470	0,237	0,530	0,293	0,144	0,252	0,826
26	2012	0,342	0,185	0,208	€ 22.300	0,464	0,236	0,536	0,300	0,142	0,257	0,825
26	2013	0,337	0,181	0,204	€ 22.000	0,456	0,235	0,544	0,309	0,134	0,263	0,832
26	2014	0,347	0,190	0,218	€ 22.400	0,445	0,239	0,556	0,317	0,128	0,272	0,833
27	2006	0,240	0,091	0,099	€ 39.300	0,269	0,472	0,731	0,259	0,129	0,264	0,81
27	2007	0,234	0,085	0,088	€ 40.400	0,265	0,471	0,735	0,264	0,134	0,264	0,811

27	2008	0,240	0,092	0,096	€ 39.800	0,262	0,470	0,739	0,269	0, 139	0,267	0,813
27	2009	0,248	0,094	0,104	€ 37.400	0,254	0,470	0,746	0,276	0,144	0,271	0,815
27	2010	0,241	0,093	0,094	€ 39.400	0,250	0,467	0,749	0,282	0,148	0,277	0,816
27	2011	0,244	0,093	0,097	€ 40.100	0,244	0,465	0,756	0,291	0,151	0,284	0,819
27	2012	0,248	0,097	0,104	€ 39.700	0,238	0,461	0,762	0,301	0,155	0,292	0,818
27	2013	0,249	0,097	0,101	€ 39.800	0,229	0,457	0,771	0,314	0,160	0,299	0,82
27	2014	0,254	0,102	0,107	€ 40.300	0,221	0,450	0,778	0,328	0,160	0,306	0,823
28	2006	0,325	0,166	0,172	€ 30.000	0,275	0,449	0,725	0,276	0,095	0,240	0,795
28	2007	0,326	0,169	0,172	€ 30.500	0,270	0,443	0,730	0,287	0,101	0,239	0,797
28	2008	0,339	0,186	0,182	€ 30.100	0,269	0,444	0,731	0,287	0,107	0,240	0,798
28	2009	0,324	0,167	0,169	€ 28.700	0,257	0,442	0,742	0,300	0,111	0,243	0,804
28	2010	0,329	0,171	0,173	€ 28.900	0,241	0,442	0,758	0,316	0,112	0,246	0,806
28	2011	0,330	0,173	0,176	€ 29.200	0,238	0,430	0,762	0,332	0,117	0,249	0,81
28	2012	0,313	0,155	0,157	€ 29.400	0,222	0,431	0,777	0,346	0,119	0,256	0,81
28	2013	0,302	0,141	0,146	€ 29.800	0,217	0,427	0,783	0,356	0,123	0,264	0,811
28	2014	0,316	0,156	0,159	€ 30.400	0,209	0,425	0,791	0,366	0,126	0,270	0,814

Appendix C: Test for normality for all three income inequality measures

	GINI	*Theil*	*MLD*
Skewness	-3.4315	-1.3553	-1.2277
Kurtosis	20.2546	6.9727	6.4871

Source: Stata

Appendix D: Spearman rank correlation

	Country	Year	GINI	Theil	MLD	GDP	Low	Med	Medhigh	High	Foreign	Oldage	Mortality
Country	1.0000												
Year	-0.0330	1.0000											
GINI	0.0707	0.1422	1.0000										
Theil	0.0677	0.1307	0.9962	1.0000									
MLD	0,0804	0.1547	0.9858	0.9861	1.0000								
GDP	-0.0757	-0.1162	-0.4074	-0.4028	-0.3869	1.0000							
Low	0.1003	-0.1873	0.1915	0.1913	0.2062	0.3901	1.0000						
Med	-0.0992	0.0038	-0.2845	-0.2787	-0.2807	-0.4513	-0.8445	1.0000					
Medhigh	-0.1004	0.1867	-0.1922	-0.1920	-0.2068	-0.3907	-1.0000	0.8455	1.0000				
High	-0.0463	0.2406	0.0379	0.0260	0.0271	0.4691	0.0007	-0.4163	-0.0024	1.0000			
Foreign	-0.1329	0.0372	-0.0594	0.0519	0.0674	0.5015	0.2163	-0.3981	-0.2172	0.4695	1.0000		
Oldage	-0.1539	0.3411	0.2839	0.2743	0.3108	-0.0014	0.0741	-0.0777	-0.0746	0.0030	0.0618	1.0000	
Mortality	0.1057	0.2889	-0.1035	-0.1016	-0.0882	0.6235	0.5291	-0.5738	-0.5297	0.3336	0.4879	0.2314	1.0000

Note: (i) The educational levels are shortened to Low, Medium, Medhigh and High. (ii) Foreign_country is shortened to Foreign. (iii) Oldagedependency is shortened to Oldage and Mortality rates is shortened to Mortality. Source: Stata.

Coefficient	Relation
0.00-0.19	Very weak
0.20-0.39	Weak
0.40-0.59	Moderate
0.60-0.79	Strong
0.80-1.00	Very strong

Source: Evans (1996)