Income Inequality and Economic Growth

An Empirical Investigation of the European Union 1991–2019

Albin Lagerlöf

Bachelor's Thesis - NEKH03

May 2021

Supervisor: Pontus Hansson

Department of Economics

Lund University - School of Economics and Management

Abstract

This thesis examines the effect of income inequality on economic growth within the European Union (EU) during the period 1991–2019. For the empirical investigation, the analysis includes four measures of income inequality. The data points for these measures have been collected from the World Income Inequality Database (WIID). The regression procedure is carried out using panel data with fixed effects, including 150 observations from 25 countries in the EU. The results in this thesis find no significant effect of income inequality on economic growth when controlling for a single measure of inequality. However, the results suggest significant effect on the growth rate when the inequality variables are included together. When included jointly, income inequality in bottom of the income distribution has significant negative effect on the growth rate. For the same regression procedure, the ratio of the top and bottom decimal of the income distribution has significant positive effect on the growth rate. Implications from these results suggest income inequality should be analysed with complementary measures. This would enable a more comprehensive observation of trends in income inequality as well as its potential effects on economic growth.

Keywords: Income Inequality, Income Distribution, Economic Growth, European Union, Panel Data

Table of contents

1. Introduction

1.1 Background

In 2010, the European Commission presented the proposal of *Europe 2020*. The document covered new strategies and paths for the cooperation and interaction between the member countries of the European Union (EU) (European Commission, 2010). Central for the strategy was to embrace a development characterised by "[…] smart, sustainable and inclusive growth" (European Commission, 2010, p. 8). As stated by the European Commission (2010), the future development of the European Union would not only focus on the economic growth, but also on its sustainability and inclusiveness. Also with certain focus on inclusive growth, the strategy stated education, structural reforms on labour market and reduced poverty to be important components for the development, future and unity of the EU (European Commission, 2010).

One subject that was not covered in the strategy plan explicitly was the matter of economic inequality. However, as argued by Fischer and Strauss (2021), the appearance of economic inequality has been a rising topic for discussion. That is, trends in economic inequality are close to the subjects of division and exclusion (Fischer & Strauss, 2021). It would therefore be of interest to discuss economic inequality in relation to the development characteristics emphasised by the *Europe 2020* strategy plan. The subjects of growth, development and economic inequality also raise further questions; Is there a relationship between economic inequality and growth? More specific, is economic inequality a determinant of the economic growth rate? Is this a concern for the European Union? These fields of questions are important to examine because of two reasons. First, research that focus on economic inequality will provide material that helps to describe both its appearance and potential consequences. Second, this form of study would also investigate what components that eventually drive or limit the rate of economic growth. Both of these subjects would be of interest for the European Union in terms of unity and future development.

1.2 Purpose

The purpose of this thesis is to examine the relationship between income inequality and economic growth within the European Union. Considering the description in section 1.1, the analysis aims to observe the effect of income inequality on the rate of economic growth. The ambition of this text is therefore to bring further evidence on the appearance and consequences of income inequality.

With respect to this subject, this thesis will attempt to answer the two following research questions:

1) To what extent has income inequality affected the economic growth for the member countries of the European Union during the period 1991–2019?

2) Is the effect of income inequality on economic growth for the member countries of the European Union between 1991–2019 different depending on what measure of income inequality that is observed?

1.3 Method and main results

For the empirical investigation of the two research questions, the analysis use four different measures of income inequality. These measures are the Gini-coefficient and three ratios of shares of the income distribution. The data points for these measures are gathered from the World Income Inequality Database (WIID). Furthermore, the statistical analysis perform regressions with fixed effects using panel data with 150 observations. This sample contain statistics from 25 member countries of the European Union, covering the time period 1991–2019. The first result presented in this thesis suggest no significant relationship between income inequality and economic growth when observing a single measure of inequality. More specific, when each of the four inequality measures are included separately in the regression we find no significant effect on economic growth. The second result from the analysis suggest significant effect of income inequality on economic growth when the inequality variables are included together. However, whether the estimate of the inequality variable is significant or not is determined by what inequality measure it is included together with. For these regressions, income inequality in bottom of the income distribution has significant negative effect on the growth rate. Contrary, the ratio of the top and bottom decimal of the income distribution has significant positive effect on the growth rate. Again, these results are not significant when the inequality measures are included separately. However, the estimates are significant when included with other inequality variables.

1.4 Structure

This thesis will be organised according to the following structure. Section 2 presents a literature review on the subject of inequality and growth. Section 3 describes the data that are used in this thesis. Section 4 presents the method used for the statistical analysis. This section describes the regression model that is applied, as well as the statistical preparations that have been considered for the data. Further, section 5 presents the results from the regressions. In turn, section 6 discusses the implications from the results according to theory and policy measures. Lastly, section 7 presents a conclusion of this thesis and areas for future research.

2. Previous literature

This section will account for previous literature on the relationship between income inequality and economic growth. Section 2.1 describes the trends of economic inequality within the European Union, and empirical findings with certain focus on the area. Section 2.2 presents theory on the link between inequality and economic growth. This section therefore provide a theoretical background on how income inequality can have effect on the growth rate. Section 2.3 presents a description of previous empirical literature on the subject. This cover the data framework that has been applied in empiric studies. Similarly, this section also explain how previous literature has applied different measures of inequality for analysis.

2.1 Inequality within the European Union

Trends in economic inequality within the European Union differ depending on what type of objectives that are being studied. As suggested by Salverda (2021), one approach is to study the EU member countries as one unit, which receives an aggregate level of income inequality for the EU. This form of categorization would be of interest if the EU should be compared to USA, China or similar regions (Salverda, 2021). In contrast, income inequality can be studied from a perspective within the EU. Filauro (2018) investigate the characteristics of net income inequality within the EU28[1], and highlight a difference between intra- and inter-country inequality. The author argue the level of income inequality in Europe between 2006 and 2014 to a majority could be explained by intra-country figures. Such perspective indicate that the level of income inequality in the EU to the majority depends on inequality levels within each member country (Filauro, 2018). Filauro (2018) suggest the general inequality trend, observed by the Gini-index, for the EU countries has decreased from 2006 to 2009. However, since then the Gini-coefficient has been relatively stable at approximately 0.35 (Filauro, 2018). Bonesmo Fredriksen (2012) confirm intra-country inequality is the major contribution to inequality levels in the EU. In addition, the author observe the specific trend of income inequality between the top and bottom of the income distribution. The author suggest the top decimal of the income distribution has increased in accumulated income in the EU between 1980 and 2008. Contrary, the share for the bottom decimal of the income distribution has decreased (Bonesmo Fredriksen, 2012). This would further support the reasoning from Fischer and Strauss (2021), who argue income inequality in the European Union may not only be explained by decreased income for

[1] EU28 refer to the former 28 member countries of the EU. However, since the withdrawal of the United Kingdom, the notation for all present member countries is EU27.

the bottom share of the income distribution. Instead, increasing income inequality would be dependent on increased share for those with already highest income (Fischer & Strauss, 2021). The following description would suggest income inequality is a concept that cannot be reduced to a single perspective. Instead, income inequality within the European Union appear differently depending on what area, time period or distribution that is being studied.

With respect to the European Union, literature has extended the research and focused also on the effect of income inequality on economic growth. Ezcurra (2007) observe the effect of income inequality on economic growth on regional level within EU. From a sample of 63 regions, the author suggest income inequality affect the growth rate negatively. As noted by the author, this result eventually lay foundation for future regional programs from the EU. Additional study by Kustepeli (2006), investigate the empirical relevance of the Kuznets curve within the EU. That is, level of income inequality may change with the level of development. The author find no evidence for the Kuznets curve to be present. As additionally presented by Kustepeli (2006), this result do not appear to dependent on the inclusion of additional member countries either. Certain focus has also been observed for the link between income inequality and membership in the EMU[2] (Bertola, 2010; Bouvet, 2010).

[2] Economic and Monetary Union within the European Union.

2.2 Theory

2.2.1 Saving and investments

De Janvry and Sadoulet (2016) describe income inequality could be positive for economic growth under the assumption that rich individuals save more of their income, compared to relatively poor individuals. As explained by the authors, if the rich have the possibility to invest their savings this could benefit the growth rate in the economy. That is, the investments from the rich would be assumed to enhance growth via positive externalities (de Janvry & Sadoulet, 2016). The same argument would also be found in the Harrod-Domar model, emphasising savings and investment opportunities for the rich would be positive for economic growth (Fields, 1989). Previous reasoning would be especially true if investment costs are high, or require access to high level of capital. This is argued by Aghion, Caroli and Garcia-Penalosa (1999), who suggest higher concentration of income and capital could generate higher rate of investments. Such state would be even more evident in an economy where access to credit is restricted (Aghion, Caroli & Garcia-Penalosa, 1999). Unequal distribution of income could also be positive for economic growth with respect to human capital. As suggested by Galor and Tsiddon (1997), if the rich invest in human capital the return could be beneficial for the growth rate. As further described by the authors, this return may for example be demonstrated through adaption of advanced technology that generate higher growth rate in the long run. This argument takes therefore dynamics of both the short and long run into account, as well as the notion of positive externalities (Galor & Tsiddon, 1997).

2.2.2 Credit constraints and credit market imperfections

Theoretical literature has also proposed inequality to affect growth via the presence of credit constraints and credit market imperfections. As noted by Barro (2000), with imperfections on the credit market, relatively poor individuals may have constraints on acquire credit. Such state could lead to investments not being fulfilled by those with low income (Barro, 2000). Furthermore, unequal distribution of capital could be negative for growth under the assumption of diminishing return of capital. This is suggested by Aghion, Caroli and Garcia-Penalosa (1999), who state the productivity of capital and investments on the margin is higher for relatively poor individuals, compared to relatively rich individuals. The authors suggest an unequal distribution of wealth would be negative for the growth rate, since this imply there are conditions where capital is more efficient. This condition would for example refer to investment and production possibilities that are missed (Aghion, Caroli & Garcia-Penalosa, 1999).

Banerjee and Newman (1993) argue the distribution of wealth would affect the individual decision when deciding on employment. First, due to constraints on obtain credit, the authors argue individuals with less capital are not able to advance in employment where high investments costs are required, and where the wage is relatively higher. Second, if distribution of capital is considered to be a determinant of the employment for individuals, the authors suggest an unequal distribution of capital would lead to a state in which initially relatively poor cannot increase their capital accumulation. In turn, this would affect the general distribution of wealth in the economy (Banerjee & Newman, 1993). As noted in the previous literature this theory discuss the distribution of wealth, which is not a form of distribution observed in this thesis. However, as proposed by Perotti (1996), income distribution and wealth distribution may be similar in studies referring to empirical evidence.

2.2.3 Fiscal policy

Inequality could also affect the growth rate via voting procedure and government policies. As discussed by Perotti (1996), this perspective apply the behaviour of the median voter, and suggest high level of inequality will indicate that the median voter is relatively poorer compared to the average voter. The median voter will thus support government policies of redistribution, which is assumed to reduce growth by the effects of markets distortions (Perotti, 1996; Neves & Silva, 2014). Therefore, the channel suggest the growth rate to be higher in economies where level of inequality is relatively low (Aghion, Caroli & Garcia-Penalosa, 1999). The theory implies that increased taxation is a result of unequal distribution of income. This has been suggested by both Meltzer and Richard (1981) and Milanovic (2000). Therefore more unequal societies consequently have a higher level of redistribution (Milanovic, 2000). Both theoretical and empirical literature, including Alesina and Rodrik (1994) and Persson and Tabellini (1994), have observed the complete concept of the fiscal policy theory. This literature therefore extends the work of Meltzer and Richard (1981) and Milanovic (2000), who does not explicitly observe the effect of inequality and redistribution on growth. Both Alesina and Rodrik (1994) and Persson and Tabellini (1994) remark the interest of conflict in economies and societies where inequality levels are high. Both studies suggest this is demonstrated by the voting procedure for redistribution policies, from which the growth rate is affected negatively. The latter is confirmed by the market distortions of taxation on, for example, level of capital (Alesina & Rodrik, 1994).

However, Bénabou (2000) argue relatively rich individuals holds a political influence as consequence of their wealth. Therefore, the support and implementation of redistribution policies would be lower in societies where wealth is unequally distributed by its population (Bénabou, 2000). Additionally, redistribution policies do not necessarily need to indicate lower growth if it leads to higher government expenditures, as suggested by Saint-Paul and Verdier (1993). For example, if the taxation process eventually fund government expenditures on education, this would be assumed to increase the level of human capital and in turn increase the growth rate (Saint-Paul & Verdier, 1993).

2.2.4 Social and political instability

Theory has also presented socio-politic unrest as link between inequality and growth. As noted by Perotti (1996), this link suggest inequality lead to higher level of socio-politic instability among both individuals and government institutions. As suggested by the author, this would negatively affect the conditions for investment and other growth enhancing activities. Therefore higher level of inequality is assumed to affect the growth rate negatively (Perotti, 1996). The effects of income inequality on growth have been studied by Venieris and Gupta (1986), who observe the effect of socio-political instability on the rate of savings. The authors find instability, expressed as demonstrations, regime and number of deaths resulted by socio-political insecurity to reduce the rate of savings. Also Alesina and Perotti (1996) use an index of social instability, however addressing larger set of variables. The authors suggest inequality leads to higher socio-politic instability, which in turn would negatively affect the level of investments. The latter, could for example be manifested by decreased level of order, or reduced holdings of property rights (Alesina & Perotti, 1996). The effects of income inequality on growth have also been studied by Keefer and Knack (2002), who argue a polarized society, defined as inequality in income, would negatively affect the function of government and policy implementation. The authors argue the higher the instability in society, the less stabile is the government and its administration. These effects eventually lower the growth rate by reduced rights of agreements and property (Keefer & Knack, 2002).

2.3 Empiric literature

2.3.1 Data setup

For the empiric literature studying economic inequality and growth, a first distinction refer to the choice between cross-section or panel data. For the studies using cross-section data, the results suggest inequality to affect the growth rate negatively (Alesina & Rodrik, 1994; Persson & Tabellini, 1994; Deininger & Squire, 1998). However, the result differ when considering specific disparities, such as regional and political differences. Persson and Tabellini (1994) find negative relationship in their cross-section analysis be present in democracies, however receiving insignificant result when observing non-democracies. Contrary, Alesina and Rodrik (1994) find no difference when including a dummy variable for democracies, but instead negative relationship in all countries included in their sample. Additionally, Clarke (1995) test for both regional and political differences and still find the relationship between inequality and growth to be negative. In terms of political system, both Clarke (1995) and Alesina and Rodrik (1994) therefore suggest the effect of inequality on growth to be negative, and not different depending on political system. Also Deininger and Squire (1998) control for regional differences and include a dummy variable for region (Latin America, Africa and Asia). The authors suggest the effect of inequality on growth still being negative, but with reduced level of significance.

In contrast, the usage of panel data present a positive relationship between inequality and economic growth (Li & Zou, 1998; Forbes, 2000; Deininger & Olinto, 2000). Forbes (2000) argue estimation using panel instead of cross-section data has the advantage of controlling for specific dimensions regarding country and time. As noted by the author, studies using cross-section data might have problems concerning omitted variables bias, which implies the coefficient of inequality variable might be misleading. Such improved statistical tool help to investigate the relationship of inequality on growth more comprehensively (Forbes, 2000). As pointed out by Li and Zou (1998), the ability to include dynamic changes may also be a more accurate description of the changes in the data for inequality and growth, compared to the framework used in cross-section data. Similar reasoning is presented by Deininger and Olinto (2000), who emphasise the shortcomings of cross-section studies when controlling for regional characteristics. This points out the selection of either cross-country or panel data, and the statistical instruments accordingly, would be of importance when studying the effect of inequality on growth.

In addition, works using panel data have also found that the result may depend on state of development. For example, Barro (2000) and Castelló-Climent (2010) apply panel data framework and find the relationship between income inequality and economic growth to be positive in developed countries. In contrast, the relationship is negative when observing less developed countries (Barro, 2000; Castelló-Climent, 2010). This would suggest conditions regarding state of development and level of income may also be components to the effect of inequality on growth. Further, this perspective is also similar to the inclusion of political system, as was discussed in the literature using cross-section data.

2.3.2 Measures of inequality

Previous mentioned studies have often used the Gini-coefficient as variable for inequality. However, the literature has also included further measures of inequality to eventually acquire new perspectives on the subject. In addition to the Gini-coefficient, Clarke (1995) include the Theil's index, the coefficient of variation, as well as ratio between the bottom 40 share and top 20 share of the country's income distribution. However, even when using different measures of inequality the author suggest the effect to be negative on the growth rate. In contrast, Persson and Tabellini (1994) measure inequality by the size of the middle class by using the middle quintile (Q3) from the income distribution. As suggested by the authors, this quintile has the property of covering the median income level. If this quintile increases, it is a sign of higher level of income for the middle class (Persson & Tabellini, 1994). Thus, this study suggest a different setup in measurement of inequality compared to previous studies. Barro (2000) use the five quintiles (Q1 to Q5) for the income distribution. However, adding the quintiles do not change the implication from using the Gini-coefficient in his study. Still the effect of income inequality on economic growth is suggested to be positive in developed countries and negative for less developed countries (Barro, 2000).

In the usage of different quintiles, Barro (2000) therefore observe the importance of different parts of the income distribution in contrast to Persson and Tabellini (1994). However, neither of the two studies observe inequality changes between certain groups. That is, neither Barro (2000) or Persson and Tabellini (1994) observe ratios, or similar measures, that eventually identify income changes between parts in the income distribution pattern. The importance of ratios has instead been emphasised by Voitchovsky (2005). As argued by the author, this attempt to answer whether income inequality may affect growth differently, when controlling for changes in different sections of the income distribution. In addition to the Gini-coefficient, Voitchovsky (2005) use the ratio between the 90th and 75th percentile, as well as the ratio between 50th and 10th percentile of the income distribution. The result from Voitchovsky (2005) suggest income inequality in top of the distribution is positive for economic growth, while income inequality in the bottom of the distribution instead is negative for economic growth. As discussed by the author, this would suggest inequality in different parts of the income distribution to have different effect on the growth rate. From this view, the findings of Voitchovsky (2005) also extend the understanding of potential relationship between income inequality and economic growth to be dependent on several measures of inequality.

3. Data

The following section describes the data used in this thesis. Section 3.1 presents the data framework applied for the analysis. Section 3.2 presents the variables for income inequality. Section 3.3 describes the dependent variable used for the model. Section 3.4 describes the additional control variables that are included for the analysis. Finally, section 3.5 provides a table including a summary of descriptive statistics for the variables that are used in this text.

3.1 Data framework

For the analysis in this text, the aim has been to gather data for all member countries in the European Union.[3] However, the selection of countries included in the sample has been with consideration to data availability. The EU member countries Croatia, Cyprus and Malta have been excluded from the sample due to limited data on income inequality. Additionally, whether the country is included in the sample considering its membership in the European Union have been with respect to the categorization in the World Income Inequality Database (UNU-WIDER, 2021a), from which the data on income inequality is gathered. Therefore, no certain consideration has been made for the year of entrance in the EU for each country. Final note refer to the inclusion of the United Kingdom, which has been included in the sample although the country is not a present member of the EU. This has been decided with regards to two aspects. First, the United Kingdom has been a member country of the EU throughout the time period that is studied. Second, adding the United Kingdom to the sample eventually increase the number of observations for the analysis.

The time frame considered in this thesis is the period 1991–2019. For the regression model, this period is computed as averages of five years.[4] Therefore, each observation represent an average of the five-year period that is being studied. This holds for all but the last period in the model. For the last period instead, the average is computed for four years. This has been done with respect to limited data for year 2020. Further description on the specification of the model for the analysis is presented in section 4.1.

[3] The countries included for analysis are Austria, Belgium, Bulgaria, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Ireland, Italy, Latvia, Lithuania, Luxembourg, Netherlands, Poland, Portugal, Romania, Slovakia, Slovenia, Spain, Sweden and the United Kingdom.

[4] The intervals have been calculated as following: 1991–1995, 1996–2000, 2001–2005, 2006–2010, 2011–2015, 2016–2019.

3.2 Variables for income inequality

The aim of this thesis is to investigate the relationship between income inequality and growth. For this to be done, the analysis will use four different measures of inequality. These are selected in order to answer both of the two research questions for this thesis. Additionally, the four measures enables to investigate the relationship between income inequality and growth more comprehensively. The first part in this section will account for the database from which the inequality measures have been gathered. The next part will describe each of the four measures of inequality that are included in the analysis.

3.2.1 World income inequality database (WIID)

All data for the inequality variables has been gathered from the World Income Inequality Database (WIID) (version 31 March, 2021), which is provided by United Nations University – World Institute for Development Economics Research (UNU-WIDER, 2021a). The WIID consists of data for various measures of inequality. Several of these are also given for the same country, during same year. What differ the data from one another is the distribution they are gathered from. Such differences may depend on resource, scale, reference unit, population or database of different quality (UNU-WIDER, 2021b). This variety of data brings possibilities for comparison. However, it also make room for measurement error. As argued by both Atkinson and Brandolini (2009) and Fields (1994), measures from different distributions may result in different value for the level of inequality. It is therefore of importance to have a defined, coherent concept of what type of inequality that is observed and used for the analysis (Atkinson & Brandolini, 2009; Fields, 1994). Also Deininger and Squire (1996) raise important aspects when measuring variables of inequality. They argue, if data for one inequality measure is gathered from different resources, this could lead to larger deviation in the inequality variable than is the actual case within each country. Second, they suggest using different distributions for inequality measures might hinder the ability to correctly compare levels of inequality also between countries. This concern may for example appear if data is gathered from a distribution that would differ significantly between countries (Deininger & Squire, 1996). For these reasons, the selection of data for the inequality measures in this thesis has been done with respect to specific criteria. Following paragraph will discuss the selection procedure for the inequality variables.

With regards to the data for inequality, the ambition for this thesis has been to select consistent and high-quality measures of income inequality. For this to be done, the following steps have been implemented in the selection of data.[5] These steps are similar to the criteria also used and implemented by UNU-WIDER (Gradin, 2021). First, the data has been gathered from net income. That is, the income level for an individual after taxation. As suggested by Atkinson and Brandolini (2009), when studying the effect of inequality on growth different forms of income may be of interest with respect to certain theoretical approaches. For example, the theory of fiscal policy is suggested to observe the distribution of gross income for the effect of inequality on growth (Atkinson & Brandolini, 2009). However, the decision of selecting net income for this study has again been with regards to data availability. Data distributed from net income has been considered to gain most consistent series for the countries in the sample. Second, the data has been prioritized from per capita scale. This means the net income has been divided by the size of each household (Gradin, 2021). Third, the data has been strived to be gathered from high quality sources. The quality indicator from UNU-WIDER (2021b) is set based on certain indicators, for example referring to coverage, type of method or consistency in the data. Applying these standards, UNU-WIDER mark a level of quality for each source (UNU-WIDER, 2021b). The WIID provide data from several sources, including LIS (Luxembourg Income Study), Eurostat, the World Bank and United Nations. In those cases when previous criteria have been met and multiple choices of source still have been available for each year, data from LIS (Luxembourg Income Study) has been selected. LIS is gathered from microdata surveys, and is by UNU-WIDER (2021b) acknowledged as one of the more comprehensive sources for income distribution data. If no data from LIS has been available, data from the source with highest quality has been selected.

Previous paragraph introduced the procedure of data selection for this thesis. As discussed, the procedure was followed with respect to potential measurement error that may occur for inequality measures. However, throughout the process there is a compromise concerning long, consistent time series and the level of quality in the data (Gradin, 2021). Therefore a final remark would be that the criteria above have been difficult to meet for some observations. That is especially with regards to countries where inequality measures are not available according to the time period analysed in this study. This was the reason why Croatia, Cyprus and Malta eventually were excluded from the sample. However, when data has been selected without

[5] The following procedure has been implemented for all four inequality variables used in this thesis.

meeting previously mentioned criteria, it has been done with respect to acquiring longer and consistent time series for the countries included in the sample. A list is provided in the beginning of Appendix, consisting of notes for those moments where data points have not been able to meet the criteria above. Further discussion on the appearance of measurement error will be provided in Section 6. However, if not otherwise stated, all data is considered to meet the selection principles described in previous paragraph.

Also other measures of inequality could have been used for the analysis. Such measures could for example have been the Theil's index, or other ratios of shares which eventually target the concept of economic inequality differently (de Janvry & Sadoulet, 2016). The four measures included in this thesis have been selected with consideration to two aspects. First, the measures have been considered to be the most available in data. Second, the measures have been selected with regards to their different and complementary properties. The usage of several measures therefore contribute to better understand, and cover, certain movements in the appearance of income inequality. The latter has also been suggested by de Janvry and Sadoulet (2016). Therefore, using the Gini-coefficient, as well as the ratios of shares, is an attempt to observe changes for different parts of the income distribution in the economy, as also advised by Odedokun and Round (2004). This thesis will argue the different measures selected, and the different changes in income inequality each cover, will be important to understand, analyse and discuss the effect of income inequality on economic growth.

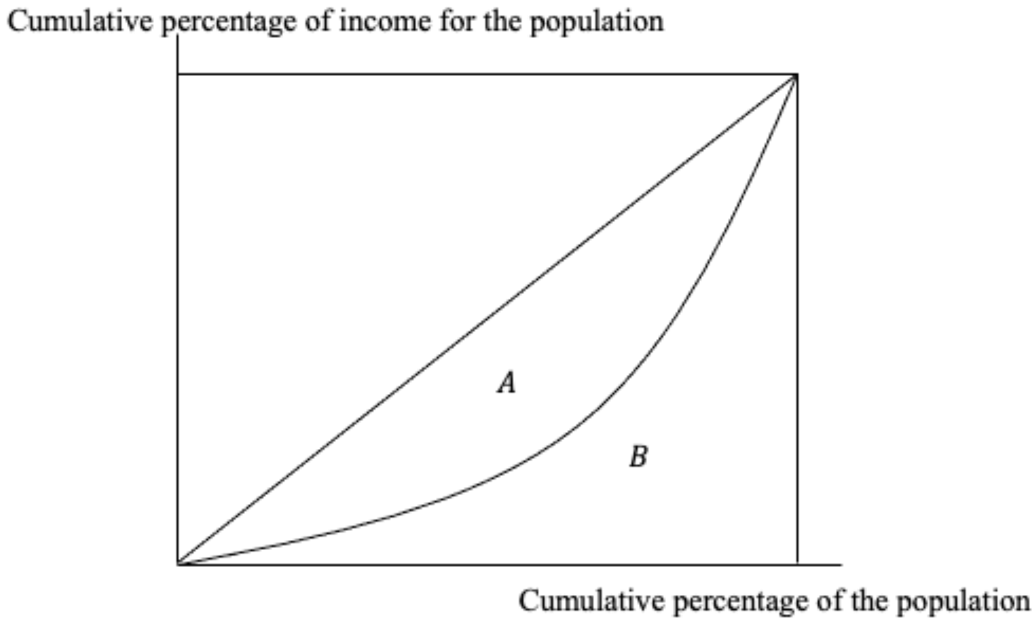
3.2.2 Gini-coefficient

The first variable used to measure income inequality is the Gini-coefficient. As previously mentioned, the data for this variable has been gathered from WIID (UNU-WIDER, 2021a). As described by de Janvry and Sadoulet (2016) is the Gini-coefficient originally derived from the Lorentz curve. This curve is graphically illustrated in graph 1 below. De Janvry and Sadoulet (2016) describe the y axis in the graph to represent the income ranking, while the x axis represent the population. The authors describe further the straight line covering the graph illustrate the state in which all resources are distributed equally. In addition, the non-linear curve found under the line is named the Lorentz curve and illustrates graphically the distribution of income in the economy (de Janvry & Sadoulet, 2016). Again noted by de Janvry and Sadoulet (2016), the Gini-coefficient can be derived from this graph according to formula (1) below, which generate a coefficient with value between 0 and 1. A Gini-coefficient with value 0 means perfect equal distribution of income, while the value 1 means perfect unequal distribution of income (de Janvry & Sadoulet, 2016). The Gini-coefficient is therefore a measure that eventually represent the general level of inequality in the economy.

$$Gini = \frac{A}{A+B} \qquad (1)$$

$$where\ 0 \leq Gini \leq 1$$

Graph 1



Source: Re-written from de Janvry and Sadoulet (2016, p. 250).

The Gini-coefficient is commonly available in data, and often recognized as measure for inequality in empiric literature (Das & Parikh, 1982; Deininger & Squire, 1996). As noted by Banerjee and Duflo (2003) and Kaplow (2005), using the coefficient therefore makes it possible for comparison between studies on the subject of economic inequality and growth. De Janvry and Sadoulet (2016) suggest the coefficient having the advantage of focusing on middle-income distribution. The authors argue, however, this could as well be a shortcoming since the coefficient may fall short for observing distribution changes in the top or bottom in the distribution pattern. Same argumentation is presented by Deininger and Squire (1996), who suggest the usage of the Gini-coefficient should be complemented by other measures. To better capture possible income changes in certain parts of the income distribution, the inclusion of quantile shares may be appropriate (Deininger & Squire, 1996).

3.2.3 Ratios

This thesis also use three ratios as measures for inequality. The three ratios are Q3Q1, Q5Q3 and P90P10. These have been recalculated using data from the WIID (UNU-WIDER, 2021a). In the database, quantiles are interpreted by dividing the population into five groups, each representing 20 percent of the population based on their income (UNU-WIDER, 2021b). The first quantile, Q1, corresponds to the 20 percent of the population with lowest income. The third quantile, Q3, refer to the middle quintile, and thus represent the middle income of the ranked distribution. The fifth quantile, Q5, refer to the 20 percent of the income distribution with the highest income. Therefore, Q3Q1 is the ratio of income between the third and first quintiles. Correspondingly, Q5Q3 is the ratio of income between the fifth and third quintiles. Each of the two ratios have been calculated by dividing the top quintile with the bottom quintile; Q3 with Q1 and Q5 with Q3, respectively. The selection of quintile ratios have been used in panel regression also by Voitchovsky (2005), as suggested measures for inequality changes in the bottom and top of the income distribution. Finally, the third ratio (P90P10) is the ratio between the top 10 percentile and the bottom 10 percentile of the income distribution. This ratio refer to the same concept as the Q3Q1 and Q5Q3 ratio. However, instead of representing shares of 20 percent, this ratio represent decimal groups. Thus the ratio conceptualize changes in distribution in the very top and bottom of the income distribution pattern (UNU-WIDER, 2021b). The P90P10 ratio has been calculated by dividing the top tenth percentile with the bottom first percentile in the income distribution ranking. With regards to the analysis in this thesis, a higher value of the ratio imply a higher spread between the two shares. In other words, a higher value of the ratio indicate higher level of income inequality. This holds for all three ratios that are included. Additionally, this is the same interpretation as for the Gini-coefficient.

3.3 Dependent variable

The dependent variable for the regression will be the average annual growth rate of Real GDP per capita. In the model, the GDP per capita growth rate will be considered to represent the growth rate for each country. Further, this form of variable is also used in previous literature on the subject of income inequality and economic growth (Perotti, 1996; Forbes, 2000). The variable has been gathered from the Penn World Table (version 10.0) by using Output-side real GDP, which in turn is expressed in purchasing power parity US dollar (year 2017) (Feenstra, Inklaar & Timmer, 2015). However, the variable has been recalculated since the Real GDP is not expressed in terms of per capita in the Penn World Table. The calculation has been made by using an additional variable for population, also this gathered from Penn World Table (version 10.0) (Feenstra, Inklaar & Timmer, 2015). Thus, GDP per capita is given by dividing the Real GDP by the population. Further, the Real GDP per capita has been calculated as annual growth rate by using formula (2) below. The start value corresponds to the GDP per capita level for the first year in each five-year period. The end value corresponds to the GDP per capita level for the last year in each five-year period. The term *n* indicate amount of years for the period that is studied. Since the model using averages of five years, this is equal to five for all but the last period. Again, due to data availability the last period is instead computed for four years. Therefore *n* equal to four in the last period.

$$\left(\frac{end\ value}{start\ value}\right)^{1/n} - 1 \qquad (2)$$

3.4 Control variables

The following section will account for the control variables included in the regression. These are variables that are assumed to be determinants of the economic growth within each country. To include them in the regression is therefore a way for the analysis to better target the potential effect of income inequality on growth, as also argued by Perotti (1996). These are also selected with background in previous literature on the subject. For each variable the predicted effect on the economic growth will be stated.

3.4.1 Human capital

The first control variable used in the model is a variable for human capital, and is gathered from the Penn World Table (version 10.0) (Feenstra, Inklaar & Timmer, 2015). The variable has been selected in order to account for the potential effects of the changes in human capital on economic growth. Previous empiric literature has used different variables for human capital. That is for example number of years in secondary school for the population of both genders separately (Perotti, 1996), or number schooling years for the adult population (Voitchovsky, 2005). The human capital variable used in this thesis is expressed in index form. The index value for each country is based on return to education, as well as average amount of years in the same (Feenstra, Inklaar & Timmer, 2015). Further, this variable is assumed to have positive effect on the growth rate. That is, increased level of human capital may, for example, enable the implementation of new technology (Barro, 1991). The variable will be named *Hum.capital* in the regression output table.

3.4.2 Population growth rate

The second control variable used in the regression is a variable for population growth. The variable is gathered from the Penn World Table (version 10.0) (Feenstra, Inklaar & Timmer, 2015). To account for the average annual growth rate for the population, also this variable has been recalculated according to formula (2) above.[6] Therefore, the variable indicate the average annual growth rate of the population. De Janvry and Sadoulet (2016) suggest population growth

[6] The decision to include four years instead of five for the last time period was mainly with respect to the variables Population growth, and Economic growth previously introduced. The calculation of formula (2) require a value for all *n*. Therefore, if data points for year 2020 are not available, the calculation need to be reduced to four years instead. If the decision had been to strictly have five year intervals for all periods, including the last, the sample size would had been reduced to a smaller set of observations. Thus, the decision of using four years for the last period is decided with respect to data availability, as also stated in section 3.1.

may have either positive or negative effect on the economic growth. As noted by the authors, population growth may in one hand be positive for economic growth as it would increase the supply of labour. On the other hand, it may have negative effect on the economic growth rate since increased population size could reduce accessibility for public goods (de Janvry & Sadoulet, 2016). The variable can therefore indicate either a positive or negative effect in the regression output table. However, a variable for population growth has been used in previous empiric literature as determinant of economic growth (Kormendi & Meguire, 1985; Levine & Renelt, 1992). The variable is therefore included in the regression considering its theoretical and empirical relevance for the model, although no predicted sign is given. In the regression output table the variable for population growth rate will be named *Pop.growth*.

3.4.3 PPP

The third control variable used in the regression is a variable for price deviation. The data for this variable is gathered from the Penn World Table (version 10.0), and refer to deviation in price level for capital formation within each country (Feenstra, Inklaar & Timmer, 2015). Therefore, the larger deviation in the price level, the larger deviation in the variable. Similar proxies for this variable have been used in previous literature as determinant of economic growth (Clarke, 1995; Alesina & Perotti, 1996; Perotti, 1996). As argued by Barro (1991), deviation in price level would have distortionary effects on investments and business activities. Therefore the variable is assumed to have negative effect on economic growth in the regression. In the regression output table, the variable will be named *PPP*.

3.4.4 Investment

Additionally, a variable for investment will be included in the regression. Data for the variable has been gathered from the World Bank (2021). The variable is referring to the Gross fixed capital formation, which is expressed as percentage of the country's GDP (World Bank, 2021). Therefore, the variable control for the level of investments within each country. Similarly as previous control variables, also investment has been used as determinant of economic growth in regression models (Levine & Renelt, 1992; Voitchovsky, 2005). Correspondingly to the results in the previous studies, the variable is assumed to have positive effect on the economic growth also in this analysis. In the regression output table the investment variable will be named *Investment*.

3.4.5 Initial GDP per capita

Last control variable included in the regression is initial GDP per capita. The variable is gathered from the same data set as the dependent variable for the regression. The variable for initial GDP per capita is therefore also expressed in purchasing power parity US dollar (year 2017) and gathered from the Penn World Table (version 10.0) (Feenstra, Inklaar & Timmer, 2015). Compared to the other variables in this regression, the variable is not calculated as an average of five years. Instead the variable represent the initial level of GDP per capita in the first year of each five-year period that is being observed. The term is used in the regression in order to control for convergence. As described by Barro & Sala-i-Martin (1992), the notion of convergence refer to endogenous growth theory, and the growth rate of different economies. More specific, the theory assume the growth rate to be higher in initially poor countries compared to the initially rich countries (Barro & Sala-i-Martin, 1992). Thus, the sign of this variable would depend conditionally on the state of economic development for the countries included in the sample. However, for the sample of 25 countries used in this analysis, almost all are categorized as 'high income' in the WIID (UNU-WIDER, 2021a).[7] With respect to development process and the sample of countries that are used, the effect of initial GDP per capita is therefore predicted to be negative on the economic growth. In the regression output table the variable for initial GDP per capita will be named *GDPpc*.

3.5 Descriptive statistics

This section presents a table of descriptive statistics for the variables used in the analysis. As given in Table 1 below, the number of observations are 150 for the 25 countries included in the sample. For each variable, also a value for mean and standard deviation from mean is given. The last two columns refer to the minimum and maximum value for each variable. With focus on the inequality variables, we can note the mean of the Gini-coefficient is 0.315 for the 25 EU countries between 1991–2019. Additionally, the mean ratio between middle and bottom quintile is 2.28, and 2.317 for the ratio of top and middle quintile. For the sample used in this text, this suggest that the middle quintile on average have 2.28 times higher income than the first quintile. Further, the top quintile have on average 2.32 times higher income than the middle quintile. In terms of difference between the ratios, this suggest a marginally higher difference in mean between the top and middle quintile, compared to the ratio of middle and bottom

[7] The only country that is not recognized as 'High income' is Bulgaria, for which the categorization is 'Upper middle income' instead.

quintile. The mean ratio between the top and bottom decimal is 9.546. Again, with respect to the countries and time period considered in this text, the mean value of P90P10 indicate that the top 10 percent on average have 9.55 times higher income, compared to the bottom 10 percent of the income distribution. Worth notice is also the higher volatility in the P90P10 variable compared to the other two ratios. This would imply income inequality observed for the very top and bottom of the income distribution eventually have higher level of variation, compared to the two quintile ratios of Q3Q1 and Q5Q3 .

Table 1

Variable	Obs.	Mean	Std. Dev.	Min	Max
Growth	150	.022	.023	-.085	.094
Gini	150	.315	.042	.202	.398
Q3Q1	150	2.28	.37	1.576	3.448
Q5Q3	150	2.317	.277	1.682	2.926
P90P10	150	9.546	3.702	3.726	26.774
Human capital	150	3.128	.334	2.02	3.808
Population growth	150	.116	.565	-1.534	1.764
PPP	150	.651	.195	.265	1.806
Investment	150	22.151	3.759	11.862	33.464
GDPpc	150	29971.79	14260.37	7174.89	86895.76

4. Method

This section presents the method that is applied for the analysis in this thesis. Section 4.1 introduces the model that will be used for the regression. Further, section 4.2 to 4.6 describes a set of statistical preparations for the data. These are performed with respect to the appropriate statistical technique for the regression.

4.1 Model specification

$$\begin{aligned} Growth_{i,t} = \beta_0 + \beta_1 Inequality_{i,t} + \beta_2 Hum.capital_{i,t} + \beta_3 Pop.growth_{i,t} \\ + \beta_4 PPP_{i,t} + \beta_5 Investment_{i,t} + \beta_6 GDPpc_{i,t} + \varepsilon_{i,t} \end{aligned} \tag{1}$$

The growth model that will be used for this thesis is presented in model (1) above. Since the observations gathered for the variables are computed for both cross-section and time, the data set is considered to be of panel data structure. As described in section 3.1, each observation is computed as an average of five years. Again, dependent variable for the regression is the growth rate given by variable $Growth_{i,t}$. For the independent variables, these are given on the right hand side of the model. The term $Inequality_{i,t}$, refer to the inequality variable that is included in the regression. Since four measures of inequality are used, there will initially be four regressions presented. Further, the variables *Hum.capital*, *Pop.growth*, *PPP*, *Investment* and *GDPpc* refer to the set of control variables that are assumed to have significant effect on the growth rate. Last, the $\varepsilon_{i,t}$ refer to the unobserved error term for the model. The β's in model (3) refer to the coefficient for each variable that is presented in the output table. More specific, the coefficient β_0 refer to the constant in the regression and β_1 refer to the coefficient for the inequality variable. Additionally, the coefficients β_2 to β_6 refer to the observed estimated coefficients for the control variables. For the model specification, *i* represent the given country that is observed, while *t* represent the given time-period that is observed. As noted by Dougherty (2016), if all cross-sectional data is available for all the time periods the panel data is said to be balanced. Since this apply also for the sample and model for this thesis, the panel data is considered to be balanced. Finally, all data has been collected and organised in Microsoft Excel. The software program Stata has been used for the statistical analysis.

4.2 Random and fixed effects

As described by Dougherty (2016), one advantage with working with panel data is the option to control for possible bias due to omitted variables. This form of problem could appear when the error terms for the cross-section unit in the sample is correlated to the explanatory variable, as a result of country-specific attributes that are not observed in the initial model (Dougherty, 2016). However, as suggested by Dougherty (2016) and Baum (2006), one can use either random or fixed effects, with respect to the unobserved characteristics for each country that is being studied. As noted by the authors, if the errors for the variables are considered as being random, the model will be using random effects. Contrary, if there is a correlation between the unobserved characteristics for each country and the variables, the model will be using fixed effects (Dougherty, 2016; Baum, 2006). To test whether random or fixed effects should be used, I use a Hausman test. This test investigate whether there is a significant difference between the random and fixed effects. The null hypothesis for the test is that the random effects is consistent and the more effective alternative, and thus should be the option to include (Baum, 2006). I perform a Hausman test for each of the four regressions, and reject the null hypothesis for all of them. This result suggest that the country-specific effects are correlated with the variables in the model. Therefore should all regressions be performed with fixed effects. The results from the Hausman tests are provided in Table 2 in Appendix.

4.3 Normality

Next I investigate whether the error terms of the four regressions are considered to be normally distributed or not. To test this, I perform each of the four regressions with fixed effects. I then use a Skewness/Kurtosis joint test on the error terms from each regression. The skewness and kurtosis refer to the attributes of the data in terms of its distribution pattern, and enables to decide on normality (Mills, 2014). Under the null hypothesis for the Skewness/Kurtosis joint test, the error terms are considered to be approximately normally distributed (Stata, n.d.). At a significance level of 5 percent, I conclude the error terms from the regressions using the Gini-coefficient, Q3Q1 and P90P10 are approximately normally distributed. In contrast, the error terms from the regression with Q5Q3 are not considered to be normally distributed. However, for the analysis in this text, I decide to keep the data for Q5Q3 in its original form. This has mainly been decided with respect to the appearance of potential outliers. To have the error terms from the Q5Q3 regression approximately normal under the joint test, I can remove outliers in the Q5Q3 data that eventually disturb the errors to be centred around its mean value. However, the reduction of outliers would have meant a reduced sample size. This may affect the fit in the regression negatively. Second, as explained by Wooldridge (2009), the decision of what is considered to be an outlier is not conclusive. Any practice of remove outliers may therefore be a subjective decision. This eventually raise the concern of possible biased result for the figures in the regression output table. To overcome this potential problem, I decide to keep the data for Q5Q3 in its original form. The output table from the Skewness/Kurtosis joint test is provided in Table 3 in Appendix.

4.4 Multicollinearity

Table 4 below illustrate the level of correlation between the variables for the analysis. Observations in column (1) on the horizontal axis, refer to the correlation between the dependent variable, and the independent variables, given by (2) to (10) on the vertical axis. With respect to the sample in this analysis, we can note at least one important aspect from the table. In the matrix, all four inequality variables have a relatively high level of correlation. In other words, in practice they observe trends in income inequality relatively similar. This may suggest that the multicollinearity could appear in our data, which emerge when two or more variables used in the right hand side of the model is highly correlated (Dougherty, 2016). The appearance of multicollinearity could make it difficult to distinguish the effect of the correlated variables on the dependent variable (Wooldridge, 2009).

To test this more formally I also perform a collinearity diagnostics test. Interpreting this test, I especially observe the VIF column which refer to the Variance Inflation Factor. As pointed out by Wooldridge (2009), the VIF column indicate the level of correlation among the independent variables. The author state there is no formal limit what is considered to be high level of correlation according to the VIF, although the figure 10 is one limit that is used in the literature as benchmark for awareness of multicollinearity. As suggested by the author, the test can be used to illustrate the behaviour of the data, and rather be used as complement to additional tests for correlation between variables. From the Collinearity diagnostics table, the Gini-coefficient and the Q5Q3 ratio have a value above 10. Again, this could indicate a problem of multicollinearity in the data. To handle this concern, the regression procedure will follow a stepwise regression. That is, each of the four inequality variables will be included separately in the model, as also previously discussed in section 4.1. As also pointed out by Wooldridge (2009), a stepwise regression procedure enables to distinguish the effect of the independent variables on the dependent variable. The Collinearity Diagnostics table is provided in Table 5 in Appendix.

Table 4

Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
(1) Growth	1.000									
(2) Gini	0.129	1.000								
(3) Q3Q1	0.062	0.836	1.000							
(4) Q5Q3	0.168	0.927	0.645	1.000						
(5) P90P10	0.156	0.843	0.858	0.710	1.000					
(6) Hum.capital	-0.026	-0.449	-0.342	-0.461	-0.391	1.000				
(7) Pop.growth	0.029	-0.241	-0.289	-0.220	-0.282	0.077	1.000			
(8) PPP	-0.280	-0.139	-0.239	-0.126	-0.264	0.273	0.361	1.000		
(9) Investment	0.207	-0.235	-0.167	-0.168	-0.183	0.077	0.036	-0.164	1.000	
(10) GDPpc	-0.140	-0.174	-0.203	-0.186	-0.281	0.324	0.688	0.577	-0.101	1.000

4.5 Autocorrelation

Additionally, I also test for autocorrelation in the data. As presented by Dougherty (2016), autocorrelation refer to the time-dimension in the sample, and appear when the error terms are either positively or negatively dependent on previous values. If autocorrelation is present the effect will mainly be considering the standard errors, and make them misleading in the results from the regression (Dougherty, 2016). To test for autocorrelation, I perform four Durbin-Watson tests, each including one inequality variable respectively. Again suggested by Dougherty (2016), if the test statistic is close to the number 2, this imply autocorrelation is not present in the sample. For all four tests I acquire a test-statistic approximately 2.[8] Therefore no autocorrelation is considered to be present in the data. One may solve the problem of autocorrelation by having longer time intervals for the observations in the sample (Dougherty, 2016). With respect to this thesis, the observations used for analysis cover intervals of five years. This technique could therefore be one explanation why autocorrelation do not appear.

[8] The test-statistic for each test is given by:

Regression including the Gini:
Durbin-Watson d-statistic(7, 150) = 2.008

Regression including Q5Q3:
Durbin-Watson d-statistic(7, 150) = 2.055

Regression including Q3Q1:
Durbin-Watson d-statistic(7, 150) = 2.000

Regression including P90P10
Durbin-Watson d-statistic(7, 150) = 2.002

4.6 Heteroscedasticity

Finally, I test whether my data is heteroscedastic or homoscedastic. As noted by Dougherty (2016), this means to test whether the variance among the error terms is considered to be constant or not over the observations. Should the variance not be constant will the standard errors of the estimates in the regression be misleading, and the data is said to be heteroscedastic (Dougherty, 2016). To test for potential heteroscedasticity in my data, I perform the Breusch-Pagan test. This is done for each of the four regressions that will be presented in the result section. The null hypothesis for the test imply the error terms are considered to be constant (Wooldridge, 2009). For all four tests, I fail to reject the null hypothesis. This means no signs of heteroscedasticity can be found the data. One explanation to this result may be due to the degree of homogeneity among the countries in EU. The results from the Breusch-Pagan tests are provided in Table 6 in Appendix.

5. Results

This section presents the results from the regressions, which are provided in two regression tables. Table 7 contain four regressions, where the inequality variables have been included separately. Table 8 contain six regressions where the inequality variables have been included together instead. As stated in section 4.1, the dependent variable for each regression is the average growth rate. Each regression also include the five control variables for the analysis.[9] Following the results from the Hausman tests, each regression is made using fixed effects. The figures in each table corresponds to the estimated coefficient for the variable and its t-statistics. Note that the t-statistics is given in parentheses. Moreover, both tables include asterisks which corresponds to the level of significance for the estimated coefficients. The end of each table also provide the R^2 value for the regressions.[10]

[9] For a more detailed description of the variables, see section 3.

[10] Since fixed effects are used, the R^2 that is provided refer to the R^2-within. This is for both Table 7 and Table 8.

Table 7

Variable	(1)	(2)	(3)	(4)
Gini	0.126			
	(1.33)			
Q3Q1		-0.00699		
		(-0.87)		
Q5Q3			0.0208	
			(1.63)	
P90P10				0.000914
				(1.27)
Hum.capital	0.110***	0.121***	0.112***	0.113***
	(6.62)	(7.09)	(6.91)	(6.93)
Pop.growth	-0.00161	-0.00288	-0.00215	-0.00226
	(-0.27)	(-0.47)	(-0.36)	(-0.38)
PPP	-0.0423***	-0.0409***	-0.0433***	-0.0396***
	(-3.63)	(-3.52)	(-3.71)	(-3.41)
Investment	0.00282***	0.00309***	0.00282***	0.00287***
	(4.84)	(5.22)	(4.88)	(4.97)
GDPpc	-2.36×10^{-6}***	-2.48×10^{-6}***	-2.33×10^{-6}***	-2.39×10^{-6}***
	(-6.90)	(-7.26)	(-6.80)	(-7.04)
Constant	-0.327***	-0.309***	-0.341***	-0.306***
	(-6.61)	(-6.49)	(-6.65)	(-6.44)
$\mathbf{R^2}$	0.440	0.436	0.445	0.440
Number of obs.	150	150	150	150

Dependent variable: GDP per capita growth rate
t-statistics in parentheses (…)
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Column (1) to (4) in Table 7 represent the four regressions where a variable for income inequality is included separately. As given in the table, the R^2 for regression (1) to (4) is between 0.436 (regression 2) and 0.445 (regression 3). Since the R^2 indicate the measure of fit, this would suggest approximately 44 to 45 percent of the variation in the economic growth rate could be explained by the variables in the regressions (Wooldridge, 2009).

As illustrated in Table 7, each of the four regressions include one inequality variable separately. Again, this procedure is performed with respect to observe the individual effect of each inequality measure on the growth rate. The same discussion was presented in section 4. However, neither of the regressions suggest any significant effect of income inequality on economic growth. That is, neither of the four inequality variables are significant in Table 7. This means we cannot discuss, or analyse, its potential effect on the growth rate any further. Contrary, almost all control variables appear significant in regression (1) to (4). At a significance level of 1 percent, the variable for human capital, level of price deviation (given by PPP), investment and initial level of GDP per capita have a significant effect on economic growth. The significant estimates for the control variables also confirm the predicted effect that was discussed in section 3.4. The one exception is the variable for population growth, for which no specific sign was predicted. Similarly, we cannot find population growth to have a significant effect on economic growth in Table 7. This may support the reasoning about its uncertain impact on the growth rate, at least for the model specification applied in this model.

Table 8

Variable	(1)	(2)	(3)	(4)	(5)	(6)
Gini	0.359	0.364***	0.0306	0.0822		
	(1.60)	(2.80)	(0.22)	(0.64)		
Q3Q1	-0.0410***	-0.0284**			-0.0107	-0.0272**
	(-2.97)	(-2.61)			(-1.32)	(-2.51)
Q5Q3	-0.0103		0.0178		0.0249*	
	(-0.48)		(0.96)		(1.91)	
P90P10	0.00187*			0.000497		0.00261***
	(1.81)			(0.51)		(2.68)
Hum.capital	0.124***	0.120***	0.111***	0.110***	0.119***	0.126***
	(7.42)	(7.18)	(6.66)	(6.61)	(7.00)	(7.51)
Pop.growth	-0.00398	-0.00283	-0.00201	-0.00183	-0.00306	-0.00459
	(-0.67)	(-0.48)	(-0.33)	(-0.30)	(-0.51)	(-0.77)
PPP	-0.0453***	-0.0477***	-0.0434***	-0.0412***	-0.0446***	-0.0398***
	(-3.91)	(-4.12)	(-3.70)	(-3.46)	(-3.82)	(-3.51)
Investment	0.00314***	0.00304***	0.00281***	0.00282***	0.00298***	0.00317***
	(5.43)	(5.28)	(4.80)	(4.82)	(5.06)	(5.48)
GDPpc	-2.42×10^{-6}***	-2.39×10^{-6}***	-2.32×10^{-6}***	-2.36×10^{-6}***	-2.37×10^{-6}***	-2.46×10^{-6}***
	(-7.22)	(-7.15)	(-6.76)	(-6.88)	(-6.92)	(-7.39)
Constant	-0.347***	-0.367***	-0.340***	-0.319***	-0.349***	-0.305***
	(-6.73)	(-7.24)	(-6.62)	(-6.14)	(-6.78)	(-6.57)
R^2	0.487	0.471	0.445	0.442	0.453	0.468
Number of obs.	150	150	150	150	150	150

Dependent variable: GDP per capita growth rate
t-statistics in parentheses (…)
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Since neither of the inequality variables were significant in Table 7, I test whether the inequality variables may present different results if they are included together. This is illustrated in Table 8, with column (1) to (6) referring to six different regressions. In column (1), all inequality variables are included together. In column (2) to (4), each of the ratios are included together with the Gini-coefficient. In column (5) and (6) the Q3Q1 is included with Q5Q3 and P90P10, respectively. There are eventually more ways to include the inequality variables than presented in Table 8. Therefore also other ways have been tested for comparison and analysis. For example, the inequality variables can be included two by two, or three by three. Additionally, all ratios can be included separately from the Gini-coefficient. However, the results from other regressions are similar to the results presented in the table above. The same is for the control variables.

As illustrated in Table 8, significant results appear for the inequality variables when they are included jointly. Interpreting the results from Table 8, I especially note regression (1) in which all inequality variables are included. In this regression, the variable Q3Q1 has a significant negative effect on the growth rate. Contrary, the variable P90P10 appear to have significant positive effect on the growth rate. For the Q3Q1 variable, the effect is significant at the 1 percent level. For the P90P10 variable, the effect is significant at the 10 percent level. Additionally, Q3Q1 and P90P10 are significant also for other regressions over the table. However, significant results for the Gini-coefficient and Q5Q3 are limited. For example, the Gini coefficient appear only significant when included together with the Q3Q1 in regression (2). Similarly, the Q5Q3 appear only significant when included with Q3Q1 in regression (5). As noted in Table 8, whether the inequality variable in question is significant or not depends on the other measure of inequality it is included together with. Turning to the control variables instead, these are similar as presented in Table 7. That is, the estimated effect on growth for each control variable confirm the predicted effect that was discussed in section 3.4. Again, and similar to the results in Table 7, population growth seem to have no significant effect on economic growth.

From Table 8 we can analyse the effect of income inequality on the growth rate further. To clarify how the table shall be interpreted, we calculate the estimated coefficient of Q3Q1 and P90P10 for regression (6). In this regression Q3Q1 and P90P10 appear significant at the 5 and 1 percent level respectively. For this regression, an increase with 0.37 in Q3Q1 would affect the growth rate negatively with approximately 1 percentage point.[11] For this variable, an increase with 0.37 refer to the standard deviation from its mean. Contrary, an increase with 3.702 in P90P10 would affect the growth rate positively with approximately 0.96 percentage points.[12] Again, 3.702 refer to the standard deviation from the mean for the variable P90P10.

[11] The standard deviation for the Q3Q1 variable is 0.37. The effect of the variable on growth is therefore calculated by 0.37 multiplied with the estimated coefficient in Table 8, regression (6). The estimated effect on growth corresponds to 0.37*-0.0272 = -0.010064. The growth rate is given in hundredths, therefore -0.010064 refer to approximately -1 percentage point on the growth rate.

[12] The effect of P90P10 on growth is found by multiplying its standard deviation (3.702) with its estimated coefficient (0.00261). This is the same form of calculation as for the Q3Q1. The estimated effect of P90P10 is therefore given by 3.702*0.00261 = 0.00966. Again, the growth rate is given in hundredths and the estimated effect of P90P10 on the growth rate is therefore approximately 0.97 percentage points.

The standard deviation for each variable is also provided in the descriptive statistics table in section 3.5.

Note that an increase in both Q3Q1 and P90P10 represent increased inequality. However, its effect on the growth rate is different. That is, an increase in Q3Q1 is suggested to have significant negative effect on the growth rate. An increase in the variable P90P10 instead, is suggested to have significant positive effect on the growth rate. Finally, for these results to hold it is necessary that all other things are equal (Wooldridge, 2009). That is, no increase or decrease in either of the other variables in the model. It may be possible that changes in other variables could higher or lower the effect of Q3Q1 and P90P10 on the growth rate (Voitchovsky, 2005). Similar comment should also be made on the inclusion of several inequality variables. If the inequality variables have a relatively high correlation and tend to change similarly over time, it may be difficult to observe an increase in one inequality variable and at the same time hold the other constant. This would be an important note for the interpretation and analysis of the inequality estimates in the table.

6. Discussion

Previous section presented results on the effect of income inequality on economic growth in the European Union between 1991–2019. With respect to the results in Table 7, we cannot find strong evidence for a relationship between income inequality and economic growth. That is, according to the estimates in the table, there seem to be no significant effect of income inequality on the growth rate. This holds for all four inequality variables that are used for the analysis. Turning to Table 8 instead, we find significant estimates for the different inequality variables. These results are however inconvenient. More specific, each significant estimate seem to depend on the other measure of inequality we control for. However, one result that eventually can be discussed from Table 8 is the measures of Q3Q1 and P90P10. These estimates appeared significant when all inequality variables were included in regression (1). Similarly, the two measures also appeared significant over the regressions in the table. As noted in previous section, an increase in both measures represent increased level of inequality. However, its effect on the growth rate is different. Following two paragraphs will present theoretical explanation to the different effects of Q3Q1 and P90P10 on the growth rate.

In table 8, income inequality in the bottom end of the distribution, given by variable Q3Q1, had negative effect on the growth rate. Since Q3Q1 cover the ratio of the middle and the first quintile of the income distribution, an increase in the measure suggest the income level for the middle quintile is rising in relation to the bottom quintile. Such changes may lead to individuals in bottom of the income distribution perceive their income to be decreasing in proportion to the middle quintile. As discussed in section 2.2, this state may result in increased number of protests, coups, or other events that decrease the level of social, political and economic stability within the country. In turn, this could affect the growth rate negatively via reduced security for investments or financial holdings (Venieris & Gupta, 1986; Alesina & Perotti, 1996; Keefer & Knack, 2002). Additionally, in the presence of imperfect institutions on the credit market, increased inequality in bottom of the distribution may hinder individuals from fulfil their investment demands, or to independently decide on their employment (Aghion, Caroli & Garcia-Penalosa, 1999; Banerjee & Newman, 1993). These theories would provide explanation to why bottom end inequality may affect the growth rate negatively. Additionally, it is likely the theories could be combined to explain the negative effect of Q3Q1 on the economic growth.

In contrast, P90P10 appeared to have significant positive effect on the growth rate in Table 8. Since the P90P10 is computed by the ratio of top and bottom decimal of the income distribution, an increase in the measure illustrate increased concentration of income to the top decimal. The theoretical explanation to why P90P10 appear positive for the growth rate would thus be reasoned from the positive effects of increased inequality. More specific, the growth enhancing processes in top of the distribution may have a higher positive effect on the growth rate, compared to the negative effects in the bottom of the distribution. If one part of the population increase their share of total income, this could lead to increased number of projects, or firms, that eventually affect the growth rate positively (Aghion, Caroli and Garcia-Penalosa 1999). Furthermore, the investments does not necessarily need to be in physical capital, but could also support growth if they are generated in dynamics of human capital (Galor & Tsiddon, 1997). The ratio of P90P10 could therefore be one measure of income inequality that to majority observe the positive effects on economic growth. That is, the positive effects of increased income for the top decimal may have larger impact on the growth rate in relation to the bottom decimal observed in the ratio. Consequently, this could be an explanation to why the variable P90P10 have positive effect on the growth rate in Table 8. It is also noteworthy that the effects of increasing inequality in top of the distribution initially was covered by Q5Q3. However, this variable appear only significant in regression (5) in Table 8. One could argue the effects of increased income for the top decimal, covered by P90P10, have a larger impact on the growth rate compared to increased income for top quintile covered by Q5Q3.

Previous paragraphs described the results in Table 8 and concluded different measures of income inequality have different effect on the growth rate. This discussion is eventually similar to what is found and discussed by Voitchovsky (2005). The author suggest inequality in the top of the income distribution encourage the growth rate, while inequality in the bottom of the distribution would have negative outcome on the growth rate. This would indicate only observing one measure of income inequality is not sufficient to understand and observe its potential effect on economic growth (Voitchovsky, 2005). This discussion has implications also for the European Union.

As stated by Bonesmo Fredriksen (2012) and Salverda (2021), the European Union do not have measures for handling income inequality explicitly. Instead, this is “[…] mostly under the authority of national governments […]” (Bonesmo Fredriksen, 2012, p. 18). However, as argued by Fischer and Strauss (2021), if income inequality is a determinant of prosperity, development and growth, it would be of importance for the European Union as institution to observe and confront the effects of income inequality (Fischer & Strauss, 2021). From what has been presented in this and previous section, it may be of interest for the European Union and its member countries, to implement the usage of different inequality measures. If policy makers implement only one measure of income inequality for analysis, this could lead to incorrect interpretation. As discovered in Table 8, instead several measures may be needed to observe the potential effects of income inequality on economic growth. Future approach for the EU to study relationship between income inequality and economic growth may therefore be using complementary analysis of different inequality measures. Such method from the EU would eventually cover the subject of income inequality and its effects on growth more comprehensively.

As discussed in section 2.3, state of development (Barro, 2000; Castelló-Climent, 2010) or political system (Persson & Tabellini, 1994) has been suggested as determinants of the relationship between inequality and economic growth. Considering the regression procedure for this thesis, dividing the sample into certain characteristics of development, political system or similar determinants could have led to different results.[13] These factors are however left out in the regressions, which remarks limitations of the analysis in this text. Additionally, an increased sample size could have presented other estimates in both Table 7 and Table 8. For example, the present sample size of 150 observations could have been extended if data from also other distributions had been collected. That is, for example, to gather data points from gross income, earnings or consumption statistics. Similarly, the data points could have been gathered from additional scales than per capita scale solely. Although such measures of data collection could have led to different results, this would have increased the level of measurement error in the analysis. This was also discussed in section 3.2.1.

[13] As stated in section 3.4.5, Bulgaria is the only country in the sample that is not classified to be 'High income' country. One could therefore argue state of development would not be a relevant factor for the sample used in this thesis. This holds if the development process is defined by level of income.

7. Conclusion

The purpose of this thesis has been to investigate the effect of income inequality on economic growth within the European Union between 1991–2019. Following this purpose, the analysis has investigated whether income inequality has a positive or negative effect on the growth rate. Additionally, the analysis has examined whether the effect of income inequality on economic growth is dependent on the measure of inequality that is observed. Following these standpoints, the study has included four measures of inequality in the model. The measures are the Gini-coefficient, and three ratios of shares of the income distribution. Additionally, all data for the inequality measures have been gathered from the World Income Inequality Database (WIID). The regressions have been made using a panel data set, considering observations from 25 countries in the European Union between 1991–2019. For the regression procedure, a set of statistical preparations was considered with regards to the characteristics of the data. From here, it was concluded that the regressions should be including fixed effects.

Concluding the results from the regressions performed in this text, we find no significant effect of income inequality on economic growth when observing a single measure of inequality. Contrary, when the inequality variables are included together, we observe a significant effect of income inequality on economic growth. The main support for significant relationship is found for inequality in the bottom end, which has a significant negative effect on the growth rate. Support is also found for the ratio of the top and bottom decimal of the income distribution, which has a significant positive effect on the growth rate.

The analysis in this text has focused on income inequality within countries of the European Union. As argued by Neves and Silva (2014), future research on the subject of inequality and growth may focus on certain links through which income inequality could affect the growth rate. With respect to the European Union, future research could focus on the relationship between inequality and growth on regional level. This form of study, as performed by Ezcurra (2007), could extend the understanding of regional differences and the inequality-growth relationship even further. Such area of research could eventually tell whether the effect of inequality on growth differ between districts and regional-specific conditions within the EU. Further, the analysis in this thesis has focused on the effects of income inequality. To observe the broad concept of economic inequality, future research may include also other distributions for analysis. Future work could focus on the effects on growth when considering distribution of wealth (Islam & McGillivray, 2020), owning of land (Deininger & Squire, 1998) or human capital (Castelló-Climent, 2010). These recommendations of research areas could eventually present a more comprehensive understanding of economic inequality within the European Union – both in its appearance and potential effects on the growth rate.

8. References

Aghion, P., Caroli, E., & Garcia-Penalosa, C. (1999). Inequality and Economic Growth: The perspective of the new growth theories, *Journal of Economic Literature*, vol. 32, no. 4, pp.1615–1660, Available through: LUSEM Library website https://lusem.lu.se/library [Accessed 1 May 2021]

Alesina, A., & Perotti, R. (1996). Income Distribution, Political Instability, and Investment, *European Economic Review*, vol. 40, no. 6, pp.1203–1228, Available through: LUSEM Library website https://lusem.lu.se/library [Accessed 24 April 2021]

Alesina, A., & Rodrik, D. (1994). Distributive Politics and Economic Growth, *The Quarterly Journal of Economics,* vol. 109, no. 2, pp.465–490, Available through: LUSEM Library website https://lusem.lu.se/library [Accessed 15 April 2021]

Atkinson, A.B., & Brandolini, A. (2009). On data: A case study of the evolution of income inequality across time and across countries, *Cambridge Journal of Economics*, vol. 33, no. 3, pp.381–404, Available through: LUSEM Library website https://lusem.lu.se/library [Accessed 30 March 2021]

Banerjee, A.V., & Duflo, E. (2003). Inequality and Growth: What can the data say?, *Journal of Economic Growth*, vol. 8, no. 3, pp.267–299, Available through: LUSEM Library website https://lusem.lu.se/library [Accessed 25 April 2021]

Banerjee, A.V., & Newman, A.F. (1993). Occupational Choice and the Process of Development, *Journal of Political Economy*, vol. 101, no. 2, pp.274–298, Available through: LUSEM Library website https://lusem.lu.se/library [Accessed 14 April 2021]

Barro, R.J. (1991). Economic Growth in a Cross Section of Countries, *The Quarterly Journal of Economics*, vol. 106, no. 2, pp.407–443, Available through: LUSEM Library website https://lusem.lu.se/library [Accessed 29 April 2021]

Barro, R.J. (2000). Inequality and Growth in a Panel of Countries, *Journal of Economic Growth*, vol. 5, no. 1, pp.5–32, Available through: LUSEM Library website https://lusem.lu.se/library [Accessed 23 April 2021]

Barro, R.J., & Sala-i-Martin, X. (1992). Convergence, *Journal of Political Economy*, vol. 100, no. 2, pp.223–251, Available through: LUSEM Library website https://lusem.lu.se/library [Accessed 29 April 2021]

Baum, Christopher. F. (2006). An Introduction to Modern Econometrics Using Stata, Texas: Stata Press

Bénabou, R. (2000). Unequal Societies: Income distribution and the social contract, *The American Economic Review*, vol. 90, no. 1, pp.96–129, Available through: LUSEM Library website https://lusem.lu.se/library [Accessed 20 April 2021]

Bertola, G. (2010). Inequality, Integration, and Policy: Issues and evidence from EMU, *Journal of Economic Inequality*, vol. 8, no. 3, pp.345–365, Available through: LUSEM Library website https://lusem.lu.se/library [Accessed 23 May 2021]

Bonesmo Fredriksen, K. (2012). Income Inequality in the European Union, OECD Economics Department Working Papers, no.952, OECD Publishing, Paris, Available online: https://www.oecd-ilibrary.org/economics/income-inequality-in-the-european-union_5k9bdt47q5zt-en [Accessed 20 May 2021]

Bouvet, F. (2010). EMU and the Dynamics of Regional Per Capita Income Inequality in Europe, *Journal of Economic Inequality*, vol. 8, no. 3, pp.323–344, Available through: LUSEM Library website https://lusem.lu.se/library [Accessed 23 May 2021]

Castelló-Climent, A. (2010). Inequality and Growth in Advanced Economies: An empirical investigation, *Journal of Economic Inequality*, vol. 8, no. 3, pp.293–321, Available through: LUSEM Library website https://lusem.lu.se/library [Accessed 21 April 2021]

Clarke, G.R.G. (1995). More Evidence on Income Distribution and Growth, *Journal of Development Economics*, vol. 47, no. 2, pp.403–427, Available through: LUSEM Library website https://lusem.lu.se/library [Accessed 21 April 2021]

Das, T., & Parikh, A. (1982). Decomposition of Inequality Measures and a Comparative Analysis, *Empirical Economics*, vol. 7, no. 1-2, pp.23–48, Available through: LUSEM Library website https://lusem.lu.se/library [Accessed 28 April 2021]

De Janvry, A., Sadoulet, E. (2016). Development Economics: Theory and practice, Oxon: Routledge

Deininger, K., & Olinto, P. (2000). Asset Distribution, Inequality, and Growth, Policy Research Working Paper Series, no.2375, The World Bank, Available through: https://econpapers.repec.org/paper/wbkwbrwps/2375.htm [Accessed 25 April 2021]

Deininger, K., & Squire, L. (1996). A New Data Set Measuring Income Inequality, *The World Bank Economic Review*, vol. 10, no. 3, pp.565–591, Available through: LUSEM Library website https://lusem.lu.se/library [Accessed 28 April 2021]

Deininger, K., & Squire, L. (1998). New Ways of Looking at Old Issues: Inequality and growth, *Journal of Development Economics*, vol. 57, no. 2, pp.259–287, Available through: LUSEM Library website https://lusem.lu.se/library [Accessed 22 April 2021]

Dougherty, C. (2016). Introduction to Econometrics, 5th edn, Oxford: Oxford University Press

European Commission. (2010). Europe 2020: A strategy for smart, sustainable and inclusive growth, Brussels, Available online: https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=celex:52010DC2020 [Accessed 20 May 2021]

Ezcurra, R. (2007). Is Income Inequality Harmful for Regional Growth?: Evidence from the european union, *Urban Studies*, vol 44, no. 10, pp.1953–1971, Available through: LUSEM Library website https://lusem.lu.se/library [Accessed 26 March 2021]

Feenstra, R.C., Inklaar R., & Timmer, M.P. (2015). The Next Generation of the Penn World Table, *American Economic Review*, vol. 105, no. 10, pp.3150–3182, Available online: www.ggdc.net/pwt [Accessed 10 April 2021]

Fields, G.S. (1989). Changes in Poverty and Inequality in Developing Countries, *The World Bank Research Observer*, vol. 4, no. 2, pp.167–185, Available through: LUSEM Library website https://lusem.lu.se/library [Accessed 1 May 2021]

Fields, G.S. (1994). Poverty and Income Distribution: Data for measuring poverty and inequality changes in developing countries, *Journal of Development Economics*, vol. 44, no. 1, pp.87–102, Available through: LUSEM Library website https://lusem.lu.se/library [Accessed 30 April 2021]

Filauro, S. (2018). The EU-Wide Income Distribution: Inequality levels and decompositions, European Commission, Available online: https://publications.europa.eu/resource/cellar/97058bfe-62f6-11e8-ab9c-01aa75ed71a1.0001.01/DOC_1 [Accessed 20 May 2021]

Fischer, G., & Strauss, R. (2021). Introduction: An overview of issues and trends, in Fischer, G., & Strauss, R. *Europe's Income, Wealth, Consumption, and Inequality*, [e-book] Oxford University Press, pp.1–38, Available through: LUSEM Library website https://lusem.lu.se/library [Accessed 16 May 2021]

Forbes, K.J (2000). A Reassessment of the Relationship Between Inequality and Growth, *The American Economic Review*, vol. 90, no. 4, pp.869–887, Available through: LUSEM Library website https://lusem.lu.se/library [Accessed 4 March 2021]

Galor, O., & Tsiddon, D. (1997). The Distribution of Human Capital and Economic Growth, *Journal of Economic Growth*, vol. 2, no. 1, pp.93–124, Available through: LUSEM Library website https://lusem.lu.se/library [Accessed 24 April 2021]

Gradin, C. (2021). WIID Companion (March 2021): Data selection. Wider Technical Note, 4/2021. Available online: https://www.wider.unu.edu/publication/wiid-companion-march-2021-data-selection [Accessed 30 April 2021]

Islam, R.M., & McGillivray, M. (2020). Wealth Inequality, Governance and Economic Growth, *Economic Modelling*, vol. 88, no. 1, pp.1–33, Available through: LUSEM Library website https://lusem.lu.se/library [Accessed 23 May 2021]

Journals:

Kaplow, L. (2005). Why Measure Inequality?, *Journal of Economic Inequality*, vol. 3, no. 1, pp.65–79, Available through: LUSEM Library website https://lusem.lu.se/library [Accessed 28 April 2021]

Keefer, P., & Knack, S. (2002). Polarization, Politics and Property Rights: Links between Inequality and growth, *Public Choice*, vol. 111, no. 1-2, pp.127–154, Available through: LUSEM Library website https://lusem.lu.se/library [Accessed 19 April 2021]

Kormendi, R.C., & Meguire, P.G. (1985). Macroeconomic Determinants of Growth: Cross-country evidence, *Journal of Monetary Economics*, vol. 16, no. 2, pp.141–163, Available through: LUSEM Library website https://lusem.lu.se/library [Accessed 29 April 2021]

Kustepeli, Y. (2006). Income Inequality, Growth, and the Enlargement of the European Union, *Emerging Markets Finance and Trade*, vol 42, no. 6, pp.77–88, Available through: LUSEM Library website https://lusem.lu.se/library [Accessed 23 May 2021]

Levine, R., & Renelt, D. (1992). A Sensitivity Analysis of Cross-Country Growth Regressions, *The American Economic Review*, vol. 82, no. 4, pp.942–963, Available through: LUSEM Library website https://lusem.lu.se/library [Accessed 6 April 2021]

Li, H., & Zou, H. (1998). Income Inequality is not Harmful for Growth: Theory and evidence, *Review of Development Economics*, vol. 2, no. 3, pp.318–334, Available through: LUSEM Library website https://lusem.lu.se/library [Accessed 22 April 2021]

Meltzer, A.H., & Richard, S.F. (1981). A Rational Theory of the Size of Government, *Journal of Political Economy*, vol. 89, no. 5, pp.914–927, Available through: LUSEM Library website https://lusem.lu.se/library [Accessed 19 April 2021]

Milanovic, B. (2000). The Median-Voter Hypothesis, Income Inequality, and Income redistribution: An empirical test with the required data, *European Journal of Political Economy*, vol 16. no. 3, pp.367–410, Available through: LUSEM Library website https://lusem.lu.se/library [Accessed 15 April 2021]

Mills, Terence. C. (2014). Analysing Economic Data: A concise introduction, Basingstoke: Palgrave Macmillan

Neves, P.C., & Silva, S.M.T. (2014). Inequality and Growth: Uncovering the main conclusions from the empirics, *The Journal of Development Studies*, vol. 50, no.1, pp.1–21, Available through: LUSEM Library website https://lusem.lu.se/library [Accessed 16 April 2021]

Odedokun, M.O., & Round, J.I. (2004). Determinants of Income Inequality and its Effects on Economic Growth: Evidence from african countries, *African Development Review*, vol. 16, no. 2, pp.287–327, Available through: LUSEM Library website https://lusem.lu.se/library - [Accessed 19 April 2021]

Perotti, R. (1996). Growth, Income Distribution, and Democracy: What the data say, *Journal of Economic growth* vol. 1, no. 1, pp.149–187, Available through: LUSEM Library website https://lusem.lu.se/library [Accessed 6 April 2021]

Persson, T., & Tabellini, G. (1994). Is Inequality Harmful for Growth?, *The American Economic Review*, vol. 84, no. 3, pp.600–621, Available through: LUSEM Library website https://lusem.lu.se/library [Accessed 26 March 2021]

Saint-Paul, G., & Verdier, T. (1993). Education, Democracy and Growth, *Journal of Development Economics*, vol. 42, no. 2, pp.399–407, Available through: LUSEM Library website https://lusem.lu.se/library [Accessed 20 April 2021]

Salverda, W. (2021). Can the European Union Contain and Improve Income Inequality?, in Fischer, G., & Strauss, R. *Europe's Income, Wealth, Consumption, and Inequality*, [e-book] Oxford University Press, pp.516–569, Available through: LUSEM Library website https://lusem.lu.se/library [Accessed 16 May 2021]

Stata. (n.d.). Skewness and Kurtosis Tests for Normality, Available online: https://www.stata.com/features/documentation/ [Accessed 20 May 2021]

The World Bank. (2021). Gross Fixed Capital Formation (% of GDP), Available online: https://data.worldbank.org/indicator/NE.GDI.FTOT.ZS [Accessed 8 April 2021]

UNU-WIDER. (2021a) World Income Inequality Database (WIID). Version 31 March 2021. Available through: https://www.wider.unu.edu/database/wiid [Accessed 30 April 2021]

UNU-WIDER. (2021b). World Income Inequality Database (WIID): User guide and data sources. Available online: https://www.wider.unu.edu/database/wiid [Accessed 28 April 2021]

Venieris, Y.P., & Gupta, D.K. (1986). Income Distribution and Sociopolitical Instability as Determinants of Savings: A cross-sectional model, *Journal of Political Economy*, vol. 94, no. 4, pp.873–883, Available through: LUSEM Library website https://lusem.lu.se/library - [Accessed 16 April 2021]

Voitchovsky, S. (2005). Does the Profile of Income Inequality Matter for Economic Growth?: Distinguishing between the effects of inequality in different parts of the income distribution, *Journal of Economic Growth*, vol. 10, no. 3, pp.273–296, Available through: LUSEM Library website https://lusem.lu.se/library [Accessed 29 March 2021]

Wooldridge, Jeffrey. M. (2009). Introductory Econometrics: A modern approach, 4th edn, Ohio: South-Western Cengage Learning

9. Appendix

Detailed information for observations deviating from the selection procedure

Country	Observation	Distribution
Romania	2001–2005	Gini-coefficient, Q3Q1, Q5Q3 and P90P10 are gathered from *consumption* (per capita).
Estonia	1996–2000	Q3Q1, Q5Q3 and P90P10 are gathered from net income, but have not been adjusted according to *per capita.*
Lithuania	1991–1995	Q3Q1, Q5Q3 and P90P10 are gathered from *gross income* (per capita).

Table 2

Hausman test for Random or Fixed effects estimator

Regression with inequality variable: Gini-coefficient
H0: Random effects are consistent and efficient
P-value: 0

Regression with inequality variable: Q3Q1
H0: Random effects are consistent and efficient
P-value: 0

Regression with inequality variable: Q5Q3
H0: Random effects are consistent and efficient
P-value: 0

Regression with inequality variable: P90P10
H0: Random effects are consistent and efficient
P-value: 0

Table 3

Skewness and Kurtosis tests for normally distributed error terms

Regression	Obs	Skewness	Kurtosis	----- Joint test ----- Adj Chi2(2)	Prob>Chi2
Including Gini	150	0.851	0.013	5.950	0.051
Including Q3Q1	150	0.922	0.070	3.350	0.187
Including Q5Q3	150	0.979	0.007	6.760	0.034
Including P90P10	150	0.775	0.032	4.760	0.093

Table 5

Collinearity Diagnostics

Variable	**VIF**	**SQRT VIF**	**Tolerance**	**R-Squared**
Growth	1.26	1.12	0.7950	0.2050
Gini	27.64	5.26	0.0362	0.9638
Q3Q1	7.35	2.71	0.1361	0.8639
Q5Q3	13.12	3.62	0.0762	0.9238
P90P10	5.22	2.28	0.1915	0.8085
Hum.capital	1.60	1.26	0.6259	0.3741
Pop.growth	2.35	1.53	0.4254	0.5746
PPP	1.75	1.32	0.5710	0.4290
Investment	1.21	1.10	0.8266	0.1734
GDPpc	2.98	1.73	0.3354	0.6646
Mean VIF:	6.45			

Table 6

Breusch-Pagan test for Heteroscedasticity

<u>Regression with inequality variable: Gini-coefficient</u>
H0: Constant Variance
P-value: 0.8460

<u>Regression with inequality variable: Q3Q1</u>
H0: Constant Variance
P-value: 0.4535

<u>Regression with inequality variable: Q5Q3</u>
H0: Constant Variance
P-value: 0.6631

<u>Regression with inequality variable: P90P10</u>
H0: Constant Variance
P-value: 0.7377