The effect of income inequality on economic growth

A comparative study between Europe and Latin America

Fabian Oppenheim

Bachelor's thesis

May 2023

Supervisor: Pontus Hansson

Department of Economics

Lund University - School of Economics and Management

Abstract

This bachelor thesis aims to examine the impact of income inequality on economic growth and assess whether the relationship differs between Europe and Latin America. Many studies have been performed to investigate the relationship between inequality and growth, and these have yielded a variation of results. As some research finds income inequality to positively impact growth, while others find the relationship to be negative, this study will explore whether the varying results could depend on regional differences. Most of the prominent papers on the subject were written in the 1990s. However, since the 1990s income inequality within countries has increased across the globe, necessitating the need for a more up to date study. The study employs regression analysis with panel data to empirically investigate the relationship, utilizing a sample of 35 countries spanning the period from 1980 to 2019. To explore regional differences, an interactive variable is incorporated into the analysis to estimate the extent to which the effect of income inequality on economic growth diverges between Europe and Latin America. The findings indicate that income inequality significantly impacts economic growth and that the effect differs between Europe and Latin America. However, it is noteworthy that the signs of the estimated coefficients vary depending on the specific inequality variable employed and whether lagged or non-lagged values of the inequality variable are utilized.

Keywords: Income Inequality, Economic Growth, Gini Coefficient, Panel Data, Latin America, Europe

Table of Contents

1. Introduction

In 2015 the United Nations presented a list with 17 goals for sustainable development. One of these goals - Goal number 10 - focus on reducing inequality within and amongst countries (United Nations, a). While a pattern of declining inequality amongst countries has been recognized, income inequality within countries has increased overall since the 1990s (United Nations, b).

Latin America has long been characterized by significant income disparities. Over the years, the region has also experienced notable fluctuations in economic performance and low levels of social mobility as well as high political instability. The persistence of high levels of inequality has posed challenges for sustainable development, social cohesion, and poverty reduction efforts in Latin American countries (Arreaza Coll, 2023). In contrast, income inequality in European countries has been far lower than in Latin America. However, in recent years, income inequality has risen significantly across Europe as well as in the rest of the world (Chancel et al. 2022).

While equality is often considered valuable in itself, it also holds significance as an influential factor affecting other aspects of society. In particular, the relationship between inequality and economic growth has for a long time been a topic of discussion. Many researchers have investigated the link between the two factors and many different, sometimes contradicting, theories have been proposed. The relationship between inequality and growth is complex and it is likely that the various channels through which the two could affect each other differ depending on the countries being studied. Given the divergent economic and sociopolitical states of Europe and Latin America, these two regions could offer a good foundation for comparison.

The purpose of this paper is to provide additional evidence regarding the impact of income inequality and economic growth. The main ambition is to investigate if the effect of inequality on growth differs between Europe and Latin America. However, in order to examine whether the effect differs, it must first be established that there is a significant effect. With respect to this, the ambition of the thesis will be to answer the following two research questions.

- Does income inequality affect economic growth?
- Does the effect of income inequality on economic growth differ between Europe and Latin America?

A regression analysis is performed to empirically investigate the two research questions. The empirical model consists of panel data which includes 19 European countries and 16 Latin American countries, with a time span ranging from 1980 to 2019. Two measurements are used in the study as variables for income inequality. The first is the Gini coefficient and the second is a ratio between the incomes of the top 10 percent of the income distribution and the bottom 50 percent. Data for income inequality is gathered from World Inequality Database (WID.world, a).

The paper will consist of the following sections. A theoretical framework based on previous research is presented in section 2. Section 3 gives a detailed description of the data and variables included in the study. The empirical model is introduced in section 4 along with the specification tests performed. Section 5 presents the results from the regression analysis while section 6 brings a detailed discussion of the estimated results. The findings of the study are summarized in section 7.

2. Previous Literature

2.1 Income Inequality and Economic Growth

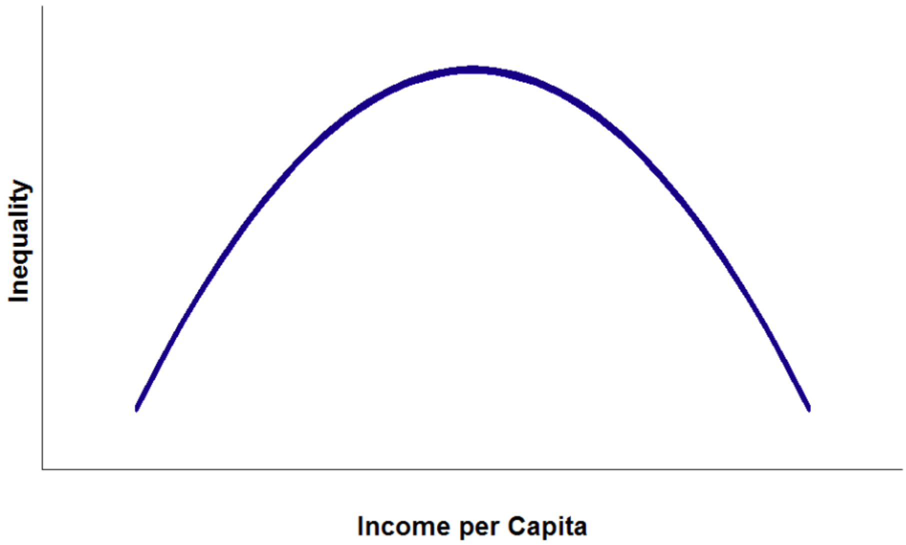
The relationship between income inequality and growth is frequently discussed in the literature on economic growth. There are many different but also contradicting views. Some researchers find that income inequality has a positive effect on economic growth while others find the relationship to be negative. It has also been investigated whether the relationship depends on what part of the income distribution is being measured. An additional important note is that the relationship is not necessarily going in one direction only. While many theories aim to explain the effect that inequality has on growth, others analyze the effect of economic growth on inequality. Since the purpose of this paper is to investigate the effect of inequality on growth, the theories and literature that will be introduced are mainly of this character. However, to provide an overall picture, some theories analyzing the effect of growth on income inequality are also included. With respect to this, the following section presents a selection of theories that seek to explain the relationship between income inequality and economic growth.

2.1.1 Traditional views and the Kuznets curve

Traditionally, fast growth has often been associated with a rise in income inequality. According to Hansson (2023) the conventional explanation is that economic growth is often driven by innovation and increased productivity in certain sectors. Initially, when productivity increases in a few sectors, only the individuals who are active in these and who have the human capital necessary to use the latest technology will benefit from the increased productivity. Eventually, as a country becomes richer, growth spreads to more sectors, resulting in a more equitable distribution of income (Hansson 2023, 39).

This relationship between income inequality and GDP per capita can be illustrated by the Kuznets curve in figure 1 below. The curve originates from Simon Kuznets who in the 1950s discovered an inverted U-shaped relationship between GDP per capita and inequality (Moffatt 2019). At very low levels of income per capita, Kuznets (1955) found that the income inequality was low. As an economy develops at an early stage, a few people could gain high incomes, thus widening the income gap. Kuznets identified this as the transition phase from an agricultural to an industrialized economy. However, as the income levels continue to increase, a middle class starts to emerge. Kuznets recognized a pattern of declining income inequality in the developed countries. According to Kuznets, an explanation could be a trend in migration flows from rural areas to the cities (Kuznets, 1955). In figure 1, the pre industrialized countries would hypothetically lie where the slope of the Kuznets curve is positive, while the industrialized would be where the income inequality is at its peak. As for the developed or post industrialized countries, income inequality is hypothetically declining and consequently these would be found where the slope of the curve is negative (Mofatt, 2019).

Figure 1:



According to the Kuznets Hypothesis, pre-industrialized countries would lie on the left hand side of the curve, where inequality and income per capita is low. As they develop and average income levels increase, inequality initially rises. At a certain level of income per capita, the level of inequality starts to decline.

2.1.2 Savings and investment

Several growth models emphasize the positive relationship of savings and investment on economic growth (see e.g. Jones and Vollrath, 2013; Todaro and Smith, 2012). In the Harrod-Domar growth model, growth of GDP is determined by net national savings, national capital-output ratio and the rate of capital depreciation. Capital output ratio is the relationship between investment and growth. The higher the ratio, the lower is the marginal productivity of capital. The fundamental intuition behind this model is that in order for an economy to grow, a certain share of the GDP must be saved and invested (Todaro and Smith 2012, 112-113). According to Barro (2000), economists commonly assume that an individual's savings rate increases with their income level. If so, then the aggregate of total savings should decrease as income is redistributed from the rich to the poor. In this way, Barro continues, increased inequality could raise the level of investments and thus stimulate growth in a transitional phase (Barro, 2000). However, even if increased inequality enables more investments to take place, what type of investments are being made are also of importance. Todaro and Smith (2012) provide several arguments with the view of inequality as an obstacle to self-sustainable growth. One of the arguments is that the investment choices, if taken by a rich elite, are unproductive and do not benefit economic growth in the long run. Todaro and Smith also state that although rich individuals save a larger dollar amount, they generally save a smaller share of their income (Todaro and Smith 2012, 219-220). Thus, they contradict the view that increased inequality raises investments, but also emphasizes that not all investments enhance sustainable growth.

2.1.3 Credit Constraints and Credit-Market Imperfections

Credit-market imperfections is commonly brought up as a channel through which income inequality could affect economic growth (e.g Barro, 2000; Aghion, Caroli & Garcia-Penalosa, 1999). According to Barro (2000), these imperfections are typically associated with asymmetric information and deficient legal institutions. With an imperfect credit market, the risk of lending is larger and thus credit will be constrained. The investment opportunities will therefore be more dependent on individuals’ wealth and income. Especially poor individuals will have problems being granted loans, and might for that reason refrain from investing in human capital. Credit constraints, which mainly affect the poor, can therefore hinder productive investments to take place (Barro, 2000). Aghion, Caroli & Garcia-Penalosa (1999) argues that credit constraints and credit market imperfections can lead to an inefficient allocation of investments. They suggest that a more equal distribution of income can lead to more growth as it increases access to credit for productive investments. In contrast, when income is unequally distributed, a smaller share of the population has access to credit, which can hinder productive investments and therefore constrain growth. Even if individuals have viable investment opportunities, they are unable to implement these because they are unable to borrow or face too high interest rates (Aghion, Caroli & Garcia-Penalosa, 1999).

2.1.4 Incentives and Efforts

One of the main arguments for income inequality to positively correlate with economic growth is the potential impact it has on incentives. Voitchovsky (2005) states that when an economic environment rewards ability, it also fosters incentives for taking risks and promoting the development of new innovations. In such an environment high income mobility is expected. With the possibility to move up (or down) in the income distribution, individual efforts are induced, both in the top and the bottom end of the distribution, Voitchovsky states. However she adds that for lower income rankings, the positive incentives might be outweighed by "worker´s feelings of frustration or unfairness" (Voitchovsky, 2005).

2.1.5 Sociopolitical unrest

Barro (2000) presents another mechanism through which income inequality could affect growth negatively. He argues that high inequality motivates poor individuals to engage in crime, riots and other disruptive activities. The stability of political institutions might even be threatened by revolution if the dissatisfaction of a broad part of the population is high enough. High political instability causes uncertainty as the expected duration of laws and rules might be reduced. Property rights might also be at risk when crime rate and political instability is high, which could deter investments. Moreover, high criminal activity and antisocial behavior negatively affects growth rate as time and resources of criminals are wasted on unproductive efforts (Barro, 2000). Furthermore, Voitchovsky (2005) states that high income inequality could lead to political polarization, where both ends of the distribution tries to expropriate the other. For instance, a rich or ruling elite might prevent pro-poor investments such as education or a certain type of infrastructure, which could be beneficial for economic development. The risk for corruption and rent seeking might also be increased when polarization and political instability is high (Voitchovsky, 2005).

2.2 Differences between Europe and Latin America

With respect to the second research question - does the effect of income inequality on economic growth differ between Europe and Latin America? - section 2.2.1 will bring light to historical factors that could explain why Latin America is lagging behind Europe in terms of economic development. Section 2.2.2 brings an additional perspective on the role of institutions. This will, together with the theories discussed above, create the framework of my own hypothesis presented in section 2.2.3.

2.2.1 Historical factors

Engerman and Sokoloff (2005) argue that extreme inequalities in Latin America followed from the type of colonies established and the specialization in plantation agriculture based on slave labor. Because of the climate in the region, Latin America had comparative advantages in sugar and other lucrative crops. This particular production demanded a high quantity of workers, which was provided in the form of slave labor. As the colonies specialized in these comparative advantages, a huge amount of slaves were imported from Africa. As a result, the population consisted of a small elite of (European colonizers), a large group of slaves (around 85% of the population) and later on non-white freedmen. Over time, extreme inequalities persisted and came to affect the development of institutions, and in turn the path of development. Institutions that evolved, such as property rights and economic opportunities, were highly undemocratic and constructed to favor a small rich elite. Investments in public goods and human capital were very limited and did not favor the broad population (Engerman and Sokoloff, 2005). Factor endowments such as land, also became very unequally distributed. This obstructed economic development as it restricted opportunities for the larger part of the population, thus inhibiting competition. The institutional framework, which concentrated power to a small rich elite, hindered reforms that would benefit the broader population. Because education for a long time was limited to a small share of the population (the rich elite), the overall level of human capital was low, making it difficult for the economy to grow. (Engerman and Sokoloff, 2000). In conclusion, there is a possibility that inequalities stemming from colonization have become entrenched in the institutional framework, negatively affecting economic development through this mechanism.

2.2.2 Institutions and Entrepreneurs

The design of institutions is often assumed to play an important role for economic development in literature on economic growth (see Baumol, 1996). Above, we discussed the potential impact unequal institutions could have on the choice of investments, factor endowments and competition et cetera. As we will see, institutions can also play a role when it comes to promoting entrepreneurship. Baumol (1996) identifies different types of entrepreneurs, and emphasizes the role institutions and policies play in stimulating the right type of entrepreneurial activity. He makes a distinction between three types of entrepreneurship: productive, unproductive and destructive. Productive entrepreneurs create economic value to society by discovering and implementing innovations, improving efficiency and creating opportunity for growth. Unproductive entrepreneurs do not create significant value to society as a whole. The type of activity Baumol mainly refers to is exploiting market imperfections, seeking monopolistic advantages or manipulating regulations (different types of rent seeking et cetera). Destructive entrepreneurs actively harm the economy and society with their activities. This category of entrepreneurial activity is associated with illegal and unethical practices, such as organized crime, that disrupt social harmony and cause economic and political instability. Baumol argues that institutions, or rules of the game as he calls it, significantly impacts whether entrepreneurship in a country predominantly consists of productive, unproductive or destructive activities (Baumol, 1996).

2.2.3 Why the effect of inequality on growth could differ

With the theories above in mind, I would like to formulate my own hypothesis. I expect income inequality to have a positive effect on economic growth in Europe and a lower (and potentially negative) effect on growth in Latin America. The reason is as follows. As Engerman and Sokoloff argue, institutions in many Latin American countries are affected by a history of extreme inequality. With respect to the Engerman and Sokoloff hypothesis discussed above, I expect these institutions to result in policymaking and investment decisions that only benefit a small part of the population at the top end of the income distribution. If this is associated with limited investments in human capital for the broader population, it could increase inequality while at the same time hindering sustainable growth. Furthermore, if a large part of the political power is concentrated to a small rich elite, through corrupt institutions, reforms necessary for economic development might not be implemented as the ones with power seek to remain at their positions. On the contrary, in European countries, where institutions are more functional, with anti-corruption laws, well-established judicial systems and property rights et cetera, it is fair to assume that the mobility between different social classes is higher. I do not say this causes less income inequality, but rather suggest that a system where many people have the possibility to move up the income rankings (or social classes) and become rich is likely more beneficial for economic growth than a system where a few people remain at the top. My reasoning here is similar to the previously discussed argument in section 2.1.4, where higher income mobility is expected to improve incentives and effort. Logically, if we assume both institutional frameworks discussed result in income inequality - the first through persisting wealth of a small elite and the latter through the possibility of becoming very rich - and the latter stimulate growth better than the first, then it is plausible that income inequality has a more positive effect on economic growth in the latter (Europe) than in the first (Latin America). In addition, if the institutions in Latin America, to a larger extent than European institutions, fail both in stimulating productive entrepreneurial activities and preventing unproductive ones (discussed in section 2.2.2), this hypothesis gains further support.

Large differences between Latin America and Europe in terms of average income level is another reason for why I expect income inequality to affect economic growth negatively in Latin America but positively in Europe. More specifically, the lowest levels of incomes in Europe might be high enough to prevent some of the negative effects of bottom end inequality. This could not be assumed for Latin America where the lowest income levels are significantly lower. Top end inequality is often positively associated with economic growth, due to previously discussed mechanisms such as incentives and effort. However, as mentioned in section 2.1.4, Voitchovsky (2005) suggests that in the bottom end of the distribution, the positive effect these mechanisms have on growth could be offset by "worker's feelings of frustration and unfairness". Instead, as discussed in section 2.1.5 high inequality might lead to poor individuals becoming marginalized. These individuals could be more inclined to start engaging in criminal activity. Furthermore, as discussed in section 2.1.3, Barro (2000) suggests the problem of credit constraints especially applies to poor individuals and poorer economies in general, as the risk of lending is larger. With higher income levels and a more functional credit market, individuals at the bottom of the distribution have better chances of being granted loans.

To conclude the last paragraph, I expect that the potentially negative effects on growth associated with bottom end inequality are more applicable in Latin American countries, while the potentially positive effects are more likely to dominate in European countries. Since the average income levels are lower in Latin America and higher in Europe, when looking at both regions as a unit Latin America could roughly be considered as the bottom end share of the income distribution while Europe would be the top end.

I am aware that the institutional frameworks within the two regions differ more than how I present them. However, some generalizations are necessary to be made. As the study investigates the average difference in effect between Europe and Latin America, and not amongst certain countries, this should not be an issue.

2.4 Empirical literature

Previous empirical literature investigating the impact of inequality on economic growth has found both positive and negative results. Much seems to depend on whether the research is conducted using panel data or cross sectional data. The studies in which cross sectional data is used in a regression analysis (see e.g Alesina & Rodrik, 1994; Persson & Tabellini, 1994; Deininger & Squire, 1998) often find income inequality to negatively affect economic growth. Studies conducted using panel data tends to result in a positive relationship (see e.g Li & Zou, 1998; Forbes, 2000; Deininger & Olinto, 2000).

Barro (2000), who uses panel data, finds that income inequality tends to retard growth in poor countries, while it encourages growth in richer nations. His panel contains 84 countries during the period 1960-1995, including 20 Sub-Saharan African countries. Moreover, Barro uses an interaction term to control for the level of GDP per capita. When GDP per capita is below $2000 (1985 U.S Dollars) he finds that growth declines as income inequality increases, while if the per capita GDP is above $2000 growth increases (Barro 2000). A study of Castelló-Climent (2010) uses a slightly more updated panel than Barro, covering 102 observations during the time period 1960-2005. In contrast to Barro who controls for GDP per capita, Castelló-Climent controls for regions. She uses a Gini coefficient based on distribution of human capital as a measure of inequality as well as a Gini coefficient based on income. Overall, Castelló-Climent finds that income inequality negatively impacts economic growth, which is an interesting result as it contradicts the previous findings from the panel data studies discussed above. However, when controlling for regions her results suggest that the effect turns positive in advanced economies and Europe, while it remains negative in low and middle-income countries (Castelló-Climent, 2010).

3. Data & Variables

The following section presents the data and the variables that the empirical model is based on. A careful compromise is necessary between including a larger number of countries or extending the time span. This is important as the decision affects the amount of observations but it can also generate different results. As mentioned in section 2.4 methods relying on cross-sectional variations tend to indicate a negative relationship of income inequality on growth, while panel data often result in a positive relationship.

The main variable in the thesis - the gini coefficient - is available for many countries within Europe and Latin America between 1980-2019. In view of this, the time period of the study extends to this period. The control variables and the sample of countries are also chosen with respect to the data availability of the gini coefficient.

The data is divided into 8 different 5 year periods, and the variables are calculated as averages over each specific time period. The two growth variables - growth in GDP per capita and population growth - are calculated as growth rates (see section 3.1).

The ambition was to include all European and Latin American countries. However, the original dataset contained holes in the data. In order to retrieve a balanced panel some countries were necessary to drop from the sample. Table 1 below lists the 35 countries from the cleaned dataset. 19 of them are European while the remaining 16 are Latin American. In summary, the panel dataset is based on 35 cross-sections and 8 time periods which adds up to 280 observations.

Table 1

Europe	Latin America
Albania	*Argentina*
Belgium	*Bolivia*
Bulgaria	*Brazil*
Switzerland	*Chile*
Cyprus	*Colombia*
Germany	*Costa Rica*
Denmark	*Dominican Republic*
Spain	*Ecuador*
Finland	*Guatemala*
France	*Honduras*
United Kingdom	*Mexico*
Greece	*Panama*
Ireland	*Peru*
Iceland	*Paraguay*
Italy	*El Salvador*
Netherlands	*Uruguay*
Norway	
Portugal	
Sweden	

3.1 Dependent Variable (*Growth)*

The dependent variable (*Growth)* is the annual average growth rate in GDP per capita. The data for GDP is gathered from Penn World Table (10.0). The measurement used is output-side real GDP expressed in purchasing power parity US dollar (year 2017) (Feenstra, Inklaar & Timmer, 2015). The decision to use this particular measure of GDP is partly based on the panel structure of the regression model. The measure is suitable for a panel regression model since it is comparable between countries but also over time periods. Using purchasing power parity makes the variable cross-sectionally comparable, as it adjusts for price variation between countries. Using a GDP measure expressed in real terms will adjust for inflation which makes the variable comparable over time.

Since this variable is expressed in total amounts and not per capita it is divided by each country's population for every given year. The data for population is also gathered from Penn World Table (10.0). The growth rate is calculated as an average over 5 year periods using equation (1) below. The end value is the value of GDP per capita the last year in each 5 year period and the start value is the value in the beginning of each 5 year period. For example, the average annual growth rate of the period 1980-1985 is calculated as $(GDP\ per\ capita\ 1985\ /\ GDP\ per\ capita\ 1980)^{1/5} - 1$. The annual average growth rate for the last time period is calculated over the 4 year period 2015-2019 due to restricted data availability.

$$\textit{Equation (1):}\ \ (end\ value\ /\ start\ value)^{1/n} - 1\ .$$

3.2 Inequality Variables

Both variables for income inequality are collected from the World Inequality Database (WID.world, a). The World Inequality Database relies on a combined effort from over a hundred researchers, including several prominent names such as Thomas Piketty, Emmanuel Saez and A.B Atkinson. Two of the main advantages of this source is the magnitude of the data and the easy accessibility (WID.world, b). The income distribution data used in the sample of this study is based on pre-tax income. This choice is primarily based on data availability. Using pre-tax income instead of net income tends to result in a higher gini coefficient, since taxation typically has a redistributive effect (Barro 2000). However, the method of using income pre-tax should not be viewed as a problem but is rather something the reader should be aware of.

Furthermore, the two variables for income inequality are used with their original value but also with one and two lagged periods. The motive for using a lagged variable is that the economy might react slowly to changes in income inequality. Voitchovsky (2005) argues similarly and suggests that a 5 year lag or longer could be reasonable. Since the data in this study is computed over 5-year periods, one lag corresponds to 5 years and 2 lags correspond to 10 years.

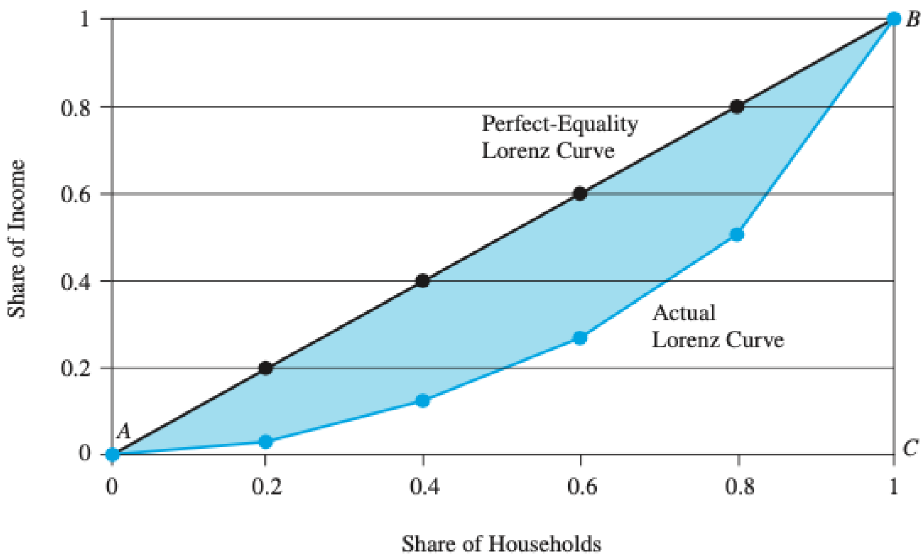
3.2.1 Gini Coefficient (*Gini)*

Perhaps the most common measure of income inequality is the Gini coefficient. This is the first variable for income inequality used in the regression model. As mentioned previously the data for this variable is collected from World Inequality Database (WID.world, a). The Gini coefficient is labeled as *Gini* in the output tables.

The gini coefficient is a general measurement of income inequality and can be derived from the Lorenz curve. The Lorenz curve shows the cumulative share of income held by a certain cumulative share of households. This is illustrated graphically by figure 1 below. Suppose we rank each household according to their income, from lowest to richest. We can then split the income groups in 5 different quintiles. The first quintile represents the poorest 20% of the population and the fifth quintile represents the richest 20%. If the income distribution is perfectly equal, then each quintile will own exactly 20 percent of the total income. This implies that the cumulative share of income will be equal to the cumulative share of the households at any point. 20% of the households will hold 20% of the total incomes, 40% of the households will hold 40% of the total incomes and so on. The actual Lorenz curve shows the share of income that each quintile actually holds. The further away the actual Lorenz curve lies from the perfect-equality Lorenz curve, the more unequal is the distribution. The gini coefficient can be calculated with equation 2 below. The value of the gini coefficient ranges between 0 and 1. A high value means that income is unequally distributed between the quintiles while a low value implies that the distribution is more equal. If the gini takes the value 0 the distribution is perfectly equal. If it takes the value 1 it means that all income is acquired by the 5th quintile (Borjas 2019, 253-254).

$$\textit{Equation 2:} \qquad Gini = \frac{blue\ area}{area\ of\ triangle\ ABC}$$

Figure 1:



The gini coefficient is only one of many ways to measure income inequality. The benefit of the gini coefficient is that it can be used as a general indicator of the inequality in a country. Data of the variable is also available for many countries for a relatively long time period. However, since variation in the gini coefficients might differ from variation in other measurements of income inequality, many researchers use other variables as complements (see e.g Voitchovsky, 2005; Deininger and Squire, 1996). This can be done in many ways. For instance, Voitchovsky (2005) uses quintile ratios to measure income inequality within the top and within the bottom of the distribution, while Deininger and Squire (1996) use the ratio of the top quintile share of income to bottom quintile share.

3.2.2 Top 10 / Bottom 50 ratio (*T10B50)*

The variable used in this study as a complement to the gini coefficient is a ratio between the income of the richest ten percent and the poorest 50 percent. The interpretation is that the higher the ratio, the bigger share holds the top 10 percent relative to the bottom 50 percent. The higher the value the more unequal is the distribution. The variable can, like the gini coefficient be considered as a general measure for inequality. However, it does capture income inequality from a different angle and I therefore consider it as a good complement. As mentioned in section 3.2, the data for this variable is collected from World Inequality Database (WID.world, a). In the regression model, the variable is referred to as *T10B50.*

3.3 Region Dummy

The main purpose of the region dummy is to produce the interactive variable (see section 3.4). However, it also allows for comparison of average growth rate between Europe and Latin America. All European countries are coded with the value 0 while the Latin American countries take the value 1. If the coefficient of the region dummy is significant the growth rate will on average differ between the two regions.[1]

3.4 Interactive Variable

A crucial part of the regression model is the interactive variable (also referred to as interaction term). This variable is used to determine whether the effect of income inequality differs between Europe and Latin America. The interactive variable is the product of the

[1] See Dougherty (2016) for the interpretation of dummy variables.

region dummy and the particular variable for income inequality in focus. Using an interactive variable is an effective way of estimating the difference in the marginal effect, of income inequality on economic growth, between the two regions. If the coefficient of the interactive variable is significant, it indicates that there is a significant difference in effect of inequality on growth, between Europe and Latin America.

However, the interpretation of an interactive variable can be complex. In order to ensure clarity, the possible outcomes are explained more thoroughly. Scenario 1, both the inequality variable and the interaction term receive statistically significant coefficients. This outcome suggests that the coefficient of the inequality variable is the effect inequality has on growth in Europe while the coefficient of the interactive variable plus or minus the coefficient of the inequality variable represents the effect of inequality on growth in Latin America. Scenario 2, the coefficient of the inequality variable is significant but the coefficient of the interactive variable is not significant. This outcome suggests that there is a significant effect of inequality on growth but the effect does not differ between the two regions. Therefore, the coefficient of the inequality variable represents the average marginal effect in both regions overall. Scenario 3, the coefficient of the interactive variable is not significant but the coefficient of the interactive variable is significant. This outcome suggests that the effect of inequality on growth in Europe is not significantly different from zero, but there is a significant effect in Latin America which corresponds to the coefficient of the interaction variable. Scenario 4, neither of the coefficients are significant. This outcome suggests that there is not a significant effect in either of the regions.[2]

3.5 Control Variables

The following section presents the control variables included in the regression model. These variables are input factors that are expected to be the main determinants of economic growth. The choice of control variables is primarily based on previous empirical literature. Levine & Renelt (1992) surveyed 41 previous growth studies and has listed the frequency in which common variables are used. The variables for the regression model are also chosen with reference to growth models in (Jones & Vollrath, 2013). Furthermore, control variables are necessary in order to prevent omitted variable bias (Nikolopoulou, 2023).

[2] See Dougherty (2016) and Frost (a) for the interpretation of interactive variables.

3.5.1 Investment *(Investment)*

Investments is the first control variable in the regression model. The data for this variable is gathered from The World Bank (2021). The variable used is gross fixed capital formation expressed as a percentage of GDP (World Bank, 2021). The variable include investments in land improvements, machinery, equipment purchases and the construction of roads, railways, schools, hospitals and buildings et cetera (World Bank, n.d). Investment share is a commonly used variable in growth regression models and is assumed to have a positive effect on economic growth (Levine & Renelt, 1992). Furthermore, capital formation is an important factor in most growth models (Jones & Vollrath 2013).

3.5.2 Population Growth *(PopGrowth)*

Population growth rate is often included as a determinant of economic growth in previous empirical literature (Levine & Renelt, 1992). It is however not perfectly clear whether population growth rate has a positive or negative effect on growth rate in GDP per capita. Population growth can lead to an increase in labor supply which could have a positive effect on growth. It also increases a country's probability of producing new ideas and therefore boosts innovation and technology. On the other hand, an increase in population means that GDP per capita will decrease ceteris paribus, which would imply a negative effect on growth in GDP per capita. In either way, the variable is an important input factor in the standard growth models (Jones & Vollrath, 2013). As mentioned previously, the data on population is gathered from Penn World Table (10.0) (Feenstra, Inklaar & Timmer, 2015). The population growth rate was calculated with the same method as the growth rate for GDP per capita (with equation 1).

3.5.3 Human Capital (*HumanCapital)*

The third control variable that will be used in the regression model is human capital. The data for human capital is collected from Penn World Table (10.0) (Feenstra, Inklaar & Timmer, 2015). Of the 41 previous growth studies surveyed by Levine & Renelt (1992) 13 included human-capital. Previous prominent studies on the subject have used different variables as a measure of human capital. Voitchovsky (2005) uses average years of schooling, while Barro (2000) uses average years of attainment of secondary and higher levels of schooling. The variable used in this study is an index based on the average years of schooling and return to education (Feenstra, Inklaar & Timmer, 2015). Return to education can be interpreted as a measure of the quality of education (Jones & Vollrath, 2013). Since human capital depends both on the time an individual spends on schooling but also the quality of the education, it can be appropriate to use an index that accounts for both aspects. However, to make the interpretation of the coefficient more intuitive, the model uses the logarithmic values of this variable. Furthermore, human capital is expected to have a positive effect on economic growth.

3.5.4 Initial GDP per capita (*GDPpc)*

The last control variable that will be used in the regression model is initial GDP per capita. The data is gathered from Penn World Table (10.0) and is based on the same measure of GDP - output-side real GDP expressed in purchasing power parity US dollar (year 2017) - as the dependent variable. The variable represents each country's per capita GDP the first year of every time period. The motive behind including this variable is to control for potential convergence within the sample. The convergence theory suggests that countries with an initially low GDP per capita grow faster than countries with a high GDP per capita, mainly due to transfer of technology (Jones & Vollrath, 2013). However this theory has met criticism by many researchers including Baumol (1986). Barro & Sala-i-Martin (1992) finds some evidence that the convergence theory holds when accounting for the fact that countries have different levels of steady state. A country that is poor relative to its steady state tends to grow faster, while a country that is rich relative to its steady state tends to grow slower. Mankiw, Romer, and Weil (1992) and Barro and Sala-i-Martin (1992) call this phenomenon conditional convergence. Furthermore, as this variable is expressed in total amounts while the others are expressed in percent, it is logarithmically transformed in order to make the interpretation of the estimated coefficient more intuitive.

3.6 Descriptive Statistics

This section presents two tables of descriptive statistics, one for each of the two regions (see table 2 and table 3 below). The European sample of countries contains 152 observations while the Latin American sample contains 128 observations.

There are several interesting notes to make from the descriptive statistics. Firstly, the annual growth rate is higher within the Latin American countries than in the European countries on average. The reported average annual growth rate is 0.023 in Europe and 0.028 in Latin

America. However, the standard deviation in the growth variable is also higher for Latin America. The reported mean of the Gini coefficient is 0.439 in Europe and 0.670 in Latin America, which suggests that the income inequality is higher in Latin America than in Europe. The standard deviation of this variable is higher in the Europe sample than the Latin American sample. The reported mean of the variable *Top 10/Bottom 50* is 7.8 in Europe and 32 in Latin America between 1980-2019. This means that in Europe on average, the richest 10 percent in a country holds 7.8 times as much income as the poorest 50 percent. In Latin America, the richest 10 percent has an income 32 times as high as the poorest 50 percent.

Table 2

Europe descriptive statistics

Variable	**Obs**	**Mean**	**Std. dev.**	**Min**	**Max**
Growth	152	.023	.025	-.033	.100
Gini	152	.439	.042	.315	.532
Top 10/Bottom 50	152	7.80	1.71	3.67	12.64
Investments	152	22.68	4.22	11.08	35.81
Pop Growth	152	.005	.006	-.0108	.023
Human Capital	152	2.47	.349	1.43	3.13
GDPpc	152	31155	14828	3681	83524

Table 3

Latin America descriptive statistics

Variable	Obs	Mean	Std. dev.	Min	Max
Growth	128	.028	.033	-.049	.143
Gini	128	.670	.038	.538	.752
Top 10 /Bottom 50	128	31.98	8.00	13.16	57.13
Investments	128	19.54	4.23	11.49	38.58
Pop Growth	128	.0164	.007	.000125	.031
Human Capital	128	1.91	.33	1.17	2.59
GDPpc	128	8677	4958	1032	24786

4. Method

This section presents the method that is applied for the study. A regression model with panel structure is used to empirically investigate the relationship between income inequality and economic growth, and whether it differs between Europe and Latin America. Panel data has several advantages over cross-sectional and time-series data. Firstly, it deals better with the problem of heterogeneity that often occurs in cross-sectional data due to omitted variable bias. Secondly, some dynamics are difficult to detect when using only cross-sectional data, but can be identified when also including the time aspect. Thirdly, it enables the model to include more observations than a model that relies solely on cross-sectional or time-series data (Dougherty 2016, 529-530). The regressions are performed in the statistical program Stata.

4.1 Model Specification

The regression model for the study is illustrated by equation 3 below.

Equation 3:

$$Growth_{i,t} = \beta_0 + \beta_1 Inequality_{i,t} + \beta_2 Investments_{i,t} + \beta_3 PopGrowth_{i,t} + \beta_4 HumanCapital_{i,t} + \beta_5 GDPpc_{i,t} + \beta_6 RegionDummy_{i,t} + \beta_7 Inequality \times RegionDummy_{i,t} + \varepsilon_{i,t}$$

It should be noted that this is the basic equation. Six variations of the regression model are estimated. *Inequality* represents income inequality and is the main variable of interest together with the interactive variable. Recall from section 3.4 that the interactive variable is the product of *Inequality* and *RegionDummy*. β_7, which is the coefficient of the interactive variable, shows how much the marginal effect of *Inequality* differs between the two regions. Depending on which inequality variable is in focus, *Inequality* refers to either *Gini, T10B50* or a lagged version of these. The other variables remain unchanged through all regressions.

4.2 Specification Tests

4.2.1 Hausman Test for Random or Fixed Effects

In order to decide whether to use a fixed effects model or a random effects model the Hausman test is performed on all regressions. The null hypothesis suggests that the difference in coefficients between the FE model and the RE model is not systematic. If the difference is not systematic (i.e H0 is true) then the random effects model should be used, as it will be more efficient than the fixed effects model. However, if there is a systematic difference in coefficients between the FE model and the RE model (i.e H0 is false), only the fixed effects model will be consistent (Dougherty, 2016). The p-values from the Hausman test are found in table 4 below.

Table 4

Main Variable in Regression Model	*Gini*	*Gini (1 lag)*	*Gini (2 lags)*	*T10B50*	*T10B50 (1 lag)*	*T10B50 (2 lags)*
P-value	0.000	0.000	0.000	0.000	0.000	0.000

As all p-values are less than 0.05 the null hypothesis is rejected for all regressions. This implies that there are significant country and time specific effects in the sample (Dougherty, 2016). The regressions will therefore be computed with the fixed effects model. Worth noting is that the fixed effects model has been used frequently in studies that investigate the relationship between inequality and growth (see e.g Benhabib and Spiegel, 1998; Forbes, 2000; Li and Zou 1998).

4.2.2 Multicollinearity

Multicollinearity is a common problem in linear regression models. A model might suffer from multicollinearity if two or more explanatory variables are highly correlated. High correlation between independent variables inflates the population variances of the estimated coefficients. This can in turn cause two problems in particular. Firstly, a high correlation can result in a lack of precision when estimating the coefficients. Secondly, it reduces the power of the model, i.e making it more difficult to identify independent variables with statistically significant effects on the dependent variable. However, high correlation between the explanatory variables is not necessarily a problem if the population variance of the error term is small, the number of observations are large and the mean square deviation of the explanatory variables are large (Dougherty, 2016).

Table 5 below presents the correlations between the variables. The lagged variables are not included as they had very similar values to the original variable but were a little bit less correlated with the other explanatory variables. Note that the correlations between the inequality variables (*Gini* and *T10B50)* and the region dummy and interactive variables (*Region_gini* and *Region_T10B50)* are very high. This is inevitable as the interaction terms are products of the inequality variables and the region dummy. The high correlation between the *Gini* and *T10B50* is not of concern, as they will not be used in the same regression model. The main focus should lie on the correlation between each inequality variable and the control variables.

Table 5

Correlation Matrix

	Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
(1)	*Growth*	1.000									
(2)	*Gini*	0.071	1.000								
(3)	*T10B50*	0.048	0.961	1.000							
(4)	*Investments*	0.104	-0.333	-0.299	1.000						
(5)	*Popgrowth*	-0.097	0.678	0.669	-0.082	1.000					
(6)	*ln_HumanCapital*	0.051	-0.646	-0.608	0.266	-0.603	1.000				
(7)	*ln_GDPpc*	-0.121	-0.731	-0.695	0.294	-0.531	0.809	1.000			
(8)	*Regiondummy*	0.081	0.944	0.909	-0.349	0.689	-0.627	-0.736	1.000		
(9)	*Region_gini*	0.078	0.957	0.935	-0.340	0.699	-0.632	-0.739	0.997	1.000	
(10)	*Region_T10B50*	0.056	0.957	0.991	-0.308	0.689	-0.614	-0.711	0.947	0.968	1.000

Each number in the top row refers to the variable with the corresponding number in the variable column.

The results presented in table 5 suggest that there is a fairly high correlation between the gini coefficient and three of the control variables (*Popgrowth, ln_hc* and *ln_GDPpc*). This is also the case for *T10B50.* However, correlation between independent variables is inevitable and is not necessarily a problem. To test whether the correlation is problematic for the regression model the "high variance inflation factors" (VIFs) are computed. The VIFs measure the extent to which multicollinearity has increased the variance of an estimated coefficient. The rule of thumb is that values above 5 are a cause for concern (Frost, b). The test results are found in table 6 below.

Table 6

Gini as main variable		*T10B50* as main variable	
Variable	VIF	Variable	VIF
Region_Gini	13.82	*Region_T10B50*	66.15
Gini	12.34	*T10B50*	59.59
ln_GDPpc	4.07	*ln_GDPpc*	3.93
ln_HumanCapital	3.47	*ln_HumanCapital*	3.47
Popgrowth	1.22	*Popgrowth*	2.40
Investments	1.22	*Investments*	1.19
Mean VIF	6.22	Mean VIF	22.79

The results from the table suggest that neither of the control variables are at major risk of causing multicollinearity. However, there is a high VIF value for both inequality variables and the interaction terms, especially in the *T10B50* ratio model. This is completely natural since the interaction terms are generated from the inequality variables (Frost, n.d). However, it could make it more difficult to find significant results, as the p-values for the estimated coefficients of the inequality variables and the interaction terms might be overestimated. It could also make the estimation of the coefficients less accurate. Therefore, the regressions are also estimated using standardized values, which according to Frost (n.d) can reduce structural multicollinearity. However, the results from the regressions with standardized values do not differ from the ones with the original values. Therefore, only the results from the original regression models are reported in section 5.

4.2.3 Normality

An assumption of the model is that the data is normally distributed. Violations of the normality assumption does not cause bias or inefficiency. However, if the sample size is small it could cause misleading results. When dealing with multiple linear regression models, the normality assumptions only apply to the residuals (Statistics Solution, 2013). Therefore, a skewness and kurtosis test is performed on the residuals from each of the six variations of the regression model. The null hypothesis states that the residuals are approximately normally distributed. If the null hypothesis is rejected, the error terms can not be considered to follow a normal distribution (Stata, n.d). All reported p-values of the test, found in table 7 below, are above the 5-percent level by a good margin. Therefore, the null hypothesis can not be rejected and we can assume the residuals are approximately normally distributed.

Table 7

				----- Joint test -----	
Main Variable in Regression	Obs	Pr(skewness)	Pr(kurtosis)	Adj chi2(2)	Prob>chi2
Gini	280	0.707	0.879	0.16	0.921
Gini (1 lag)	279	0.148	0.534	2.49	0.287
Gini (2 lags)	278	0.053	0.727	3.90	0.142
T10B50	280	0.152	0.695	2.22	0.329
T10B50 (1 lag)	279	0.124	0.6037	2.65	0.266
T10B50 (2 lags)	278	0.053	0.689	3.94	0.139

4.2.4 Autocorrelation

The model is also tested for autocorrelation. According to Dougherty (2016), autocorrelation occurs when there is correlation between error terms from different time periods. While this does not cause the OLS estimators to be biased or inconsistent, it does make OLS inefficient. It also makes the variance formula incorrect which leads to inconsistent standard errors and therefore incorrect inference. (Dougherty 2016). The Durbin-Watson test will be computed to test for autocorrelation in the data. As stated by Dougherty (2016) The Durbin-Watson test statistic lies between the values 0 and 4. The closer the reported statistic is to the value 2, the less likely is the sample to suffer from autocorrelation. If there is negative autocorrelation the value should be larger than 2 and if there is positive autocorrelation the value should be less than 2 (Dougherty 2016). Kenton (2021) states that values which lie between 1.5 and 2.5 are relatively normal and are often considered acceptable. Values outside this range could be a cause for concern (Kenton 2021). The test is computed on the six different variations of the regression model. The reported statistics are found in table 8 below.

Table 8

Main Variable in Regression Model	*Gini*	*Gini (1 lag)*	*Gini (2 lags)*	*T10B50*	*T10B50 (1 lag)*	*T10B50 (2 lags)*
Durbin-Watson Test Statistic	1.763	1.794	1.777	1.764	1.789	1.780

The values imply that there could be hints of positive autocorrelation in the sample. However, it should not be a major concern, since all the values lie inside the range of 1.5 and 2.5. Nevertheless, the p-values of the coefficients should be analyzed with some caution, as the standard errors might be slightly underestimated.

4.2.5 Heteroskedasticity

Lastly, the regression model is tested for heteroskedasticity. One of the assumptions of the model is that the data is homoscedastic, which means that the variation of the error terms are constant. If the variance of the error terms are not constant, the model suffers from heteroskedasticity. Heteroskedasticity causes the standard errors of the coefficients to be inconsistent, which results in unreliable inference and hypothesis testing. Hence, the t-statistics and p-values can not be trusted if heteroskedasticity is present (Dougherty, 2016) The Breusch-Pagan test is computed for all 6 variations of the regression model. Under the null hypothesis the data is homoskedastic. Rejecting the null implies that the data suffers from heteroskedasticity (Statology, 2020). Table 9 below presents the results from the test.

Table 9

Main Variable in Regression Model	*Gini*	*Gini (1 lag)*	*Gini (2 lags)*	*T10B50*	*T10B50 (1 lag)*	*T10B50 (2 lags)*
P-value	0.096	0.113	0.150	0.075	0.107	0.161

The reported p-values from all tests are larger than 0.05. This suggests that the data does not suffer from heteroskedasticity. However, the p-values of the models that include *Gini* and *T10B50* as main variables are fairly close to 0.05. Therefore, the results should be viewed with some caution.

5. Results

This section presents the results from the regressions. As stated in section 3.1, the dependent variable in all regressions is average growth rate in GDP per capita. The numbers in the top row indicate which of the 6 regression models is in focus. Furthermore, see section 3.4 for clarification on the interpretation of the interactive variable. With this in mind, the results from the regressions are presented in table 10 below.

Table 10

Variable	(1)	(2)	(3)	(4)	(5)	(6)
Gini	.174** (0.0858)					
Gini (1 lag)		-.0999 (0.0642)				
Gini (2 lags)			-.121** (0.0587)			
T10B50				.00399* (0.00205)		
T10B50 (1 lag)					-.00268* (0.0154)	
T10B50 (2 lags)						-.00285** (0.00142)
Interactive Variable	-.104 (0.16)	.290*** (0.1)	.248*** (0.089)	-.00371* (0.00221)	.00356** (0.0016)	.00339** (0.00146)
Investment	.00231*** (0.000468)	.00205*** (0.000448)	.00216*** (0.000448)	.00228*** (0.000468)	.00199*** (0.000448)	.0021*** (0.000449)
Popgrowth	-1.054** (0.454)	-.747* (0.451)	-.488 (0.480)	-1.05** (0.456)	-.748* (0.452)	-.516 (0.481)
ln_HumanCapital	.187*** (0.0316)	.188*** (0.0316)	.206*** (0.0321)	.187*** (0.0317)	.187*** (0.0319)	.202*** (0.0323)
ln_GDPpc	-.0511*** (0.00805)	-.0467*** (0.00771)	-.0530*** (0.00782)	-.051*** (0.00798)	-.0457*** (0.00781)	-.0513*** (0.00781)
Constant	.270*** (0.0771)	.258*** (0.0621)	.325*** (0.0587)	.311*** (0.0604)	.282*** (0.0586)	.326*** (0.0559)
R^2	0.267	0.274	0.271	0.266	0.270	0.266
OBS	280	279	278	280	279	278

Values on the right side of the variables refer to the estimated coefficients. Standard errors are found in the parentheses. The values of the coefficients and standard errors are rounded *up to three significant figures. ***, **, * indicates that the coefficient is significantly different from 0 at the 1, 5 and 10% significance levels, respectively.*

In 4 of the 6 regressions, a significant result is found for both the coefficient of the inequality variable and the coefficient of the interaction term. These results suggest that income inequality has a significant effect on economic growth and that this effect differs between Europe and Latin America. The regression outputs with significant values for both the inequality variable and the interaction variable are found in the columns (3), (4), (5) and (6). As shown in table 10, these results vary in terms of significance level. Interestingly, the signs of the inequality coefficient and the signs of the interaction term coefficient shifts, when lagged values of the inequality variable are used instead of the original variable.

As reported in column (4), the inequality variable *T10B50* without lag is found to positively affect growth in Europe at the 10% significance level. The effect is significantly lower in Latin America (at the 10 % significance level). The results in column (3) - from the regression with *Gini (2 lags)* - indicate that income inequality has a significant negative effect on economic growth in Europe and that the effect is higher in Latin America. The inequality coefficient is significantly negative at the 5% level and the coefficient of the interactive variable is significantly positive at the 1% level. Similar findings are made when looking at the values from the regressions with the lagged *T10B50* as the inequality variable. The regression with a two period lag on *T10B50* - given by column (6) - produced coefficients significantly different from zero (at the 5 % level), with a negative coefficient of the lagged *T10B50* and a positive coefficient of the interaction term. The results support the findings from the regression with the 2 period lagged *Gini,* i.e that income inequality negatively affects growth rate in Europe and that the effect is higher in Latin America than in Europe. The same results, but with lower significance, follows from the regression with a one lag version of *T10B50.* It should be mentioned that no conclusion from these regression can be drawn regarding whether the effect in Latin America is positive or negative, as the confidence intervals are too large. The sign of the coefficient will therefore only tell us whether the effect is higher or lower than in Europe.

The remaining 2 regressions - column (1) and column (2) - gave statistically significant coefficients for either the inequality variable or the interaction term (but not both). The results in column (1) suggest that income inequality - when measured as *Gini* - has a significant positive effect on economic growth, at the 5 percent significance level. However, since the coefficient of the interactive variable is not statistically significant, the effect is not found to differ between Europe and Latin America. The coefficient of the *Gini* therefore reflects the average marginal effect of inequality on growth in both Europe and Latin America. The results from the regression with the *Gini* with one lagged period, found in column (2), shows a statistically significant positive effect for the interaction term at the 1% level, but a coefficient for the *Gini* which is not statistically significant. This result suggests that there is no effect of inequality on growth in Europe, but there is a positive effect in Latin America.

Furthermore, all control variables except for *Popgrowth* are significant with the expected sign at the 1% level in all regressions. *Investment* affects economic growth positively in all regression models. Likewise, a positive coefficient is found for *ln_HumanCapital* in all regressions. Per capita income level, or *ln_GDPpc* negatively impacts economic growth. The results for *Popgrowth* suggest population growth has a negative effect on economic growth. It receives a significant coefficient at the 5% level in two regressions, at the 10% level in two regressions and non-significant results in 2 regressions. All results for the control variables are expected and in line with previous research and theories (see section 3 where the variables are presented). The *Constant* is significant at the 1 % level in all six regressions.

6. Discussion

Following from the results presented in the previous section, there are several interesting observations to make. Evidence for a relationship between income inequality and growth is found in all 6 regressions. However, the evidence varies, both in terms of appearance but also in level of significance, depending on how the regression is performed.

From the regressions where a non-lagged value of the inequality variables is used, inequality is found to have a significant positive effect on economic growth. The *Gini* is more statistically significant than the *T10B50.* As mentioned in the previous section, while the interaction term in the *T10B50* model is significant, this is not the case for the *Gini* model. We can therefore conclude that inequality, when measured as the Gini coefficient, has a positive effect on economic growth overall in Europe and Latin America. When measured as the ratio between the richest 10 percent and the poorest 50 percent, the evidence suggests that inequality has a positive effect on growth in Europe, and the effect is significantly lower (and potentially negative) in Latin America. The results from the *T10B50* regression model are similar to findings by Barro (2000). As stated in section 2.4, Barro (2000) who also uses panel data, finds that income inequality positively affects growth in rich countries, while it negatively affects growth in poor countries. In addition, the results are in line with findings of Castelló-Climent (2010), who finds that the effect is positive in Europe and advanced economies but negative in low-income countries (see section 2.4).

The results from the *T10B50* regression is the only evidence provided in support of my hypothesis presented in section 2.2.3. As can be recalled, the hypothesis suggested income inequality to have a positive effect on economic growth in Europe and that the effect would be lower (or even negative) in Latin America. Despite the coefficients being significant only at the 10% level, the findings do provide some evidence for the hypothesis and are therefore of significant importance for the study. A weakness of the study is that although the effect is found to be significantly lower in Latin America than in Europe, it can not be concluded whether the effect is negative or positive in Latin America. This is, as mentioned in section 5, because the confidence intervals of the coefficients are too large.

The reported positive effect of inequality on growth is also supported by previous empirical literature that rely on panel data. As stated in section 2.4, several studies (Li & Zou, 1998; Forbes, 2000; Deininger & Olinto, 2000) that have used panel data find income inequality to have a positive effect on economic growth. The positive relationship between income inequality and growth is also in line with some of the theories discussed in section 2.1. As stated in section 2.1.2, Barro (2000) argues that higher income inequality might affect growth positively through an increased amount of investments. In addition, as stated by Voitchovsky (2005), the positive relationship could also follow from improved incentives and efforts, which in turn follows from the possibility to move up income rankings and become rich (discussed in section 2.1.4).

The results from the regressions with lagged inequality variables, provide several interesting observations. First of all, the signs of the coefficients shift. The coefficient of both inequality variables, which were positive in the non-lagged regressions, became negative when income inequality was lagged. The coefficient of the interaction term shifts from negative to positive when a lag is introduced. The lagged regressions also yield a higher level of significance than the non-lagged regression. The strongest evidence for a relationship between inequality and economic growth is found when a two period lag is used on the inequality variables. The results suggest that income inequality has a negative effect on economic growth in Europe, and that the effect is significantly higher in Latin America (potentially positive). These findings contradict my hypothesis presented in 2.2.3, which suggested the opposite. Neither are the results similar to any of those from previous studies discussed in section 2.4. As mentioned above (and in section 2.4), Barro (2000) finds that inequality positively affects growth in rich countries while it negatively affects growth in richer places. Perhaps the most important difference between this study and the one of Barro is the sample. Barros' timespan ranges from 1960-1995 while the time period of this study contains observations between 1980-2019. His sample includes 84 countries of which 20 are from Sub-Saharan Africa. Barro also uses an interaction term to control for level of GDP per capita, while the interaction term in my regressions controls for region, which is another difference that could explain the contrasting results.

Although the result of the lagged regressions is not what I expected, theories presented in section 2 could possibly provide an explanation for this. The negative effect of inequality on growth in Europe could possibly be explained by the mechanisms which I assumed to be more applicable in Latin America. For instance, as discussed in section 2.1.4, increased incentives and effort from high income mobility is assumed to positively affect economic growth. At the bottom of the distribution these incentives might be outweighed by frustration. I expected that the positive incentives would be higher in Europe and that inequality would rather give rise to a higher degree of political instability and crime (as discussed in section 2.1.5) in Latin America. It is possible that the effect these mechanisms would have on economic growth was overestimated.

Furthermore, the contradicting results from the regressions with and without a lag on the inequality variables could be considered to weaken the evidence for a relationship between inequality and growth. On the other hand, it is possible that the relationship between the variables changes as they approach each other in time. As discussed in section 2.1.1 traditional views suggest that inequality initially increases in early stages of a growth process but then decreases again as growth spreads to more sectors. Although this procedure attempts to explain how inequality is affected by growth - while this study examines the opposite - it still indicates that the relationship between the variables might differ depending on the length of the time gap.

7. Conclusion

The purpose of this thesis has been to investigate if income inequality has an effect on economic growth in Europe and Latin America with a main focus on whether the effect differs between the two regions. Two measurements have been used as a variable for income inequality. The first is the Gini coefficient and the second is a ratio between the incomes of the top 10 percent of the income distribution and the bottom 50 percent. Data for income inequality has been gathered from World Inequality Database (WID.world, a). The relationship between income inequality and economic growth has been empirically examined with a panel data regression model using fixed effects. The sample has included a total of 280 observations from Europe and Latin America during the period 1980-2019.

Income inequality is found to have a significant effect on economic growth and the effect differs between Europe and Latin America. However, depending on whether lags are included or not both the signs of the coefficients and the levels of significance vary. First, the most statistically significant results are found when an inequality variable with a 10 year lag (2 periods) is used. This effect is found to be negative in Europe and higher (potentially positive) in Latin America than in Europe. This holds true whether income inequality is measured using the Gini coefficient or as a ratio between the incomes of the top 10 percent and the bottom 50 percent of the population. Second, when a non lagged ratio of the richest 10 percent and the poorest 50 percent is used, the effect is instead found to be positive in Europe, and lower (potentially negative) in Latin America than in Europe. The evidence for these results are however not as strong as in the case of the lagged variables. The results from the non lagged Gini coefficient suggest that income inequality positively affects growth overall in Europe and Latin America. Although it is possible that the relationship between the variables changes as they approach each other in time, the contradicting results from the regressions with and without a lag could also be considered to weaken the evidence for the relationship between inequality and growth.

Moreover, the relationship between income inequality and economic growth is complex, as evidenced by the conflicting findings in existing literature. For future research, this study could be extended to include some of the various channels through which inequality is expected to influence growth. By doing so, we could gain a deeper understanding of the connection between inequality and growth and for why the relationship differs between regions.

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. 37, no. 4, pp.1615–1660, Available through: https://www.aeaweb.org/articles?id=10.1257/jel.37.4.1615 [Accessed 2 May 2023]

Alesina, A., & Rodrik, D. (1994). “Distributive Politics and Economic Growth”, *The Quarterly Journal of Economics,* vol. 109, no. 2, pp.465–490, Available through: https://www.jstor.org/stable/2118470 [Accessed 10 May 2023]

Arreaza Coll, Adriana. (2023). “Latin America’s Inequality is Taking a Toll on Governance” *Americas Quarterly.* Available Through: https://www.americasquarterly.org/article/latin-americas-inequality-is-taking-a-toll-on-governance/ [Accessed 21 May 2023]

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 12 April 2023]

Barro, R.J., & Sala-i-Martin, X. (1992). “Convergence”, *Journal of Political Economy*, vol. 100, no. 2, pp.223–251, Available through: https://www.journals.uchicago.edu/doi/10.1086/261816 [Accessed 18 April 2023]

Baumol, W.J. (1990), “Entrepreneurship: Productive, Unproductive, and Destructive”, *Journal of Political Economy*, vol. 98, 893–921. Available Through: https://www.jstor.org/stable/2937617 [Accessed 8 May 2023]

Baumol, W.J. (1986) “Productivity Growth, Convergence and Welfare: What the Long-run Data Show.” *The American Economic Review.* Vol. 76, No. 5, pp. 1072-1085. https://www.jstor.org/stable/1816469?seq=9 [Accessed 18 April 2023]

Castelló-Climent, Amparo. (2010). “Inequality and growth in advanced economies: an empirical investigation”, *The Journal of Economic Inequality.* Available through: https://link.springer.com/article/10.1007/s10888-010-9133-4 [Accessed 20 May 2023]

Chancel, Lucas; Piketty, Thomas; Saez, Emmanuel; Zucman, Gabriel. (2022). “World Inequality Report 2022” *World Inequality Database.* Available Through: https://wir2022.wid.world/www-site/uploads/2021/12/Summary_WorldInequalityReport2022_English.pdf [Accessed 28 Mars 2023]

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: https://www.sciencedirect.com/science/article/pii/S0304387898000996 [Accessed 12 April 2023]

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 21 April 2023]

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: https://econpapers.repec.org/paper/cufwpaper/512.htm [Accessed 21 April 2023]

Dougherty, Christopher. 2016). “Introduction to Econometrics” 5th Edition. Oxford. Oxford University Press.

Feenstra, Robert C., Robert Inklaar and Marcel P. Timmer (2015), "The Next Generation of the Penn World Table" *American Economic Review*, 105(10), 3150-3182, available for download at www.ggdc.net/pwt [Accessed 28 Mars 2023]

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: https://www.aeaweb.org/articles/pdf/doi/10.1257/aer.90.4.869 [Accessed 10 May 2023]

Frost, Jim. (a). “Understanding Interaction Effects in Statistics”. *Statistics by Jim*. Available through: https://statisticsbyjim.com/regression/interaction-effects/ [Accessed 18 May 2023]

Frost, Jim. (b). “Multicollinearity in Regression Analysis: Problems Detection and Solutions.” *Statistics by Jim*. Available through: https://statisticsbyjim.com/regression/multicollinearity-in-regression-analysis/ [Accessed 18 May 2023]

George J. Borjas. (2020). “Labor Economics". 8th Edition. New York. McGraw Hill Education.

Hansson, P. (2023). “Kompletterande kompendium”: Economic Growth. Course material, Department of Economics, Lund University.

Jones, Charles I. & Dietrich Vollrath (2013), “Introduction to Economic Growth”, 3rd edition. New York. W.W. Norton & Company.

Kenneth L. Sokoloff and Stanley L. Engerman. (2000) “History lessons: Institutions, factor endowments, and paths of development in the New World,” *Journal of Economic Perspectives* 14: 217–232 Available Through: https://www.aeaweb.org/articles?id=10.1257/jep.14.3.217 [Accessed 25 April 2023]

Kenneth L. Sokoloff and Stanley L. Engerman. (2005). “Colonialism, inequality, and long-run paths of development.” *National Bureau of Economic Research*. Available Through: https://www.nber.org/papers/w11057 [Accessed 25 April 2023]

Kenton, Will. (2021). “Durbin Watson Test: What It Is in Statistics, With Examples”. Investopedia. Available through: https://www.investopedia.com/terms/d/durbin-watson-statistic.asp [Accessed 10 May 2023]

Kuznets, Simon. (1955). “Economic Growth and Income Inequality”: *The American Economic Review*, Vol. 45, No. 1, pp. 1-28 (30 pages), Available through: https://www.jstor.org/stable/1811581?seq=27 [Accessed 25 April 2023]

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: https://www.researchgate.net/publication/4901181_A_Sensitivity_Analysis_of_Cross-Country_Growth_Regressions [Accessed 5 April 2023]

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: https://econpapers.repec.org/paper/cufwpaper/74.htm [Accessed 19 April 2023]

Mankiw, N. Gregory, David Romer and David Weil. (1992) "A Contribution to the Empirics of Economic Growth." *Quarterly Journal of Economics* 107: 407-437 Available through: https://scholar.harvard.edu/mankiw/publications/contribution-empirics-economic-growth [Accessed 18 April 2023]

Michaud, Céline. (2022). "Europe weakened by rising inequality" *Voxeurop.* February 8. Available through: https://voxeurop.eu/en/europe-too-inequality-is-growing/ [Accessed 21 May 2023]

Moffatt, Mike. (2019) "Essential Economics Terms: Kuznets Curve." ThoughtCo: thoughtco.com/kuznets-curve-in-economics-1146122 [Accessed 25 April 2023]

Nikolopoulou, K. (2023). What Is Omitted Variable Bias? Definition & Examples. *Scribbr.* Available through: *https://www.scribbr.com/research-bias/omitted-variable-bias/* [Accessed 17 May 2023]

Persson, T., & Tabellini, G. (1994). Is Inequality Harmful for Growth?, *The American Economic Review*, vol. 84, no. 3, pp.600–621, Available through: https://www.jstor.org/stable/2118070 [Accessed 10 May 2023]

Statistics Solutions. (2013). "Normality". https://www.statisticssolutions.com/academic-solutions/resources/directory-of-statistical-analyses/normality/ [Accessed 11 May 2023]

Stata.com. (n.d) sktest - Skewness and Kurtosis test for normality. https://www.stata.com/manuals13/rsktest.pdf [Accessed 11 May 2023]

Statology. (2020) "The Breusch Pagan Test: Definition & Examples". https://www.statology.org/breusch-pagan-test/ [Accessed 14 May 2023]

The World Bank. (2021) Investments: *Gross fixed capital formation as percent of GDP* https://data.worldbank.org/indicator/NE.GDI.FTOT.ZS?end=2021&start=1980 [Accessed 8 April 2023]

The World Bank. (n.d). Metadata glossary. https://databank.worldbank.org/metadataglossary/world-development-indicators/series/NE.GDI.FTOT.ZS [Accessed 8 April 2023]

Todaro, Michael P., and Stephen C. Smith. "Economic Development". 11th ed., Addison-Wesley, 2012.

United Nations. (a) "The 17 Goals" Department of Economic and Social Affairs https://sdgs.un.org/goals [Accessed 21 May 2023]

United Nations. (b) "Inequality - Bridging the Divide" Department of Economic and Social Affairs https://www.un.org/en/un75/inequality-bridging-divide [Accessed 21 May 2023]

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: https://link.springer.com/article/10.1007/s10887-005-3535-3 [Accessed 28 March 2023]

WID.world. (a). World Inequality Database (WID.world). Available through:https://wid.world/data/ [Accessed 28 Mars 2023]

WID.world. (b). World Inequality Database (WID.world). Available through: https://wid.world/wid-world/ [Accessed 28 Mars 2023]