Stockholm School of Economics
Department of Economics
BE551 Degree Project in Economics
Spring 2023

The Effect of EU Funding on Income Inequality and Employment

Andres Cruz Sanchez (25199) and Daniela Dahlgren (24930)

Abstract

An outspoken goal of the European Union is to decrease differences between member states and promote cohesion. To achieve this, the European Union provides funding to the regions with greatest development needs. However, during the last decades, the within-country income inequality level has risen in the European Union. Earlier studies have mainly focused on measuring the effect of the funding on economic growth and have overlooked other important aspects. Therefore, this paper sets out to examine the effect of the funding on income inequality and employment levels within the NUTS 2 regions of the European Union. This was carried out by using a fuzzy regression discontinuity design, which exploits the scheme of the funding programme. No statistically significant results were found, which questions the efficiency of the funding.

Keywords: Income Inequality, Employment, European Union, EU Policy

JEL: D31, J60, I30, O52

Supervisor: Fabio Blasutto
Date submitted: May 14, 2023
Date examined: May 23, 2023
Discussants: Shuanghao Tang & Jonathan Stefansson
Examiner: Johanna Wallenius

Acknowledgements
First and foremost, we would like to thank Fabio Blasutto, our supervisor at the Stockholm School of Economics, for the help and guidance throughout the process of writing this thesis. We would also like to thank our friends and family for their support.

"The test of our progress is not whether we add more to the abundance of those who have much; it is whether we provide enough for those who have too little."

Franklin D. Roosevelt

Table of Contents

1. Introduction

Equality is a philosophical question; how do we define it; how do we measure it; how do we deal with it; should we deal with it? This is a matter which both great and small minds alike have thought about throughout the ages. Deng Xioping, who laid the foundations for the emergence of China as an formidable economic player on the world stage, encouraged ‘some people to get rich first’ so that others could follow in their wake (Naughton, 1993). Similarly, Nietzsche proposed that egalitarianism – the belief that people should have the same rights and opportunities since we are all equal – was inhibitive for human progress (Satz & White, 2021). In 1955, Ludwig von Mises wrote that “Inequality of wealth and incomes is the cause of the masses' well-being, not the cause of anybody's distress”. Mises was part of the liberal thinkers who believed that any redistribution efforts within society would reduce incentives and thereby, in extension, thought that economic inequality would spur economic growth (Lansley, 2012). However, today, there are other voices arguing that income inequality is a problem that should be addressed. These voices stem from organizations such as the European Union (EU) and the United Nations who argue that higher levels of income inequality may hamper economic development in the long term (European Commission, n.a.a.; United Nations, 2017)

Worldwide, the standard of living has increased substantially. However, this progress has not been equal in all areas and this has given rise to vast inequalities around the world (Stanley, 2022). In Europe, the overall income inequality has increased over the last few decades. While the differences between countries have decreased, the income inequality within countries has increased (Boschma, 2022; Papatheodorou & Pavlopoulos, 2013). Specifically looking at the EU, it has been estimated that from 1980 to 2009 the within-country income inequality has increased by approximately 50% (Doran & Jordan, 2013). Figure 1 in the appendix shows how the top 10% of income earners’ share of national income has changed. While the top earners’ share has increased in almost all parts of Europe to varying extents, the bottom 50% of income earners’ share has decreased all over the continent. Additionally, this divide between the rich and the poor within Europe is continuing to grow (Blanchet et al., 2019).

Income inequality can be problematic for a society simply because it may seem unfair and lead to discontent among its citizens. Furthermore, it also seems as income inequality is connected to many other factors that affect the functioning of society. High income inequality has been linked to distrust against politicians in Europe (Khun et al., 2014). In regions which are economically lagging behind, there is evidence of a direct mistrust against the EU rather than against the national institutions. This implies that high income inequality may lead to rising nationalism and euroscepticism in these specific regions (Schraff & Lipps, 2020). Lastly, high levels of income inequality seem to be correlated with lower levels of well-being and happiness in Europe (Delhey & Dragolov, 2014). This effect is especially strong for European low income earners (Alesina et al., 2014).

The construction of the single market in the EU has been shown to contribute to income inequality in Europe. This regional integration has further been linked to the decrease in between-country inequality and the increase in within-country inequality (Beckfield, 2006;

Beckfield, 2009). In order to correct for imbalances, such as income inequality, and to improve the economy of the EU overall, the EU has a cohesion policy which involves multiple funds whose aims differ (European Commission, 2015a). Through these funds, the EU aims at achieving cohesion among its member states, which is an objective that has become one of its key policies (Farole et al., 2011). Furthermore, the EU could be considered a special case as it is a union of multiple, self-governing countries and does not have the same kind of fiscal equalization tools as a sovereign state would have. This implies that this funding programme is a vital part of the measures that the EU can apply in order to correct imbalances. They correct these imbalances by promoting investment in research and development, education and training, among other things. They also target public expenditures in order to enhance the single market of the member countries. The cohesion policy has existed since 1989 and before each seven-year programming period, specific development targets are set (Barnier, 2003).

Thus, given the effects income inequality has on society and the cohesion policy which aims at correcting such imbalances, we believe that it is of high interest to study the effect of the funding. In this paper, we will focus on the funds that are administered through a specific threshold mechanism and that we deem to be among the most central funds (European Commission, 2007; European Commission, 2015b). Current literature trying to evaluate the impact of the funds mainly focus on economic growth. However, given the difficulties of finding a conclusive answer to the effect of the funds, we believe that it would be valuable to use other measures as well in order to better understand the overall effect of the funding.

Following the article *Going NUTS: The effect of EU Structural Funds on Regional Performance* written by Becker et al. (2010), we use a similar methodology. However, instead of focusing on economic growth, we study income inequality within the EU at the regional NUTS 2 level. Additionally, similarly to the above mentioned article, we will examine the effect on employment growth. This paper will focus on the period from 2007 to 2020, which consist of two programming periods: one period from 2007 to 2013 and the second period from 2014 to 2020. The research questions will be the following:

1. Has the funds had an impact on the income inequality within NUTS 2 regions in the European Union between the years 2007 to 2020?

2. Has the funds had an impact on the employment levels within NUTS 2 regions in the European Union between the years 2007 to 2020?

Different research methods have been used in previous studies to estimate the effect of the funds. A generalized method of moments (GMM) estimator, spatio-temporal econometric model, difference-in-difference (DID) estimator, ordinary least squares (OLS), fixed effects model (FE) and regression discontinuity design (RDD) are some research methods that have been applied in this context – the latter three being the most commonly used methods (Butkus et al., 2019; Mohl & Hagen, 2010). This paper will investigate the research questions using a fuzzy RDD in order to assess the effect. In brief, an RDD exploits a discontinuity created by some threshold value, which changes the probability of treatment (Stock & Watson, 2019). In this paper, the discontinuity that will be exploited is the threshold set up by the EU for receiving funding. This threshold is 75% of the average regional GDP/capita of the EU and NUTS 2 regions below this level are eligible (European Commission, n.a.c.; European Commission, n.a.d.).

The fuzziness in the RDD arises because some NUTS 2 regions who are eligible for funding do not receive it, whilst others that are not eligible do receive it. This is in contrast to a sharp RDD, where the eligibility criteria would have to be followed strictly. Due to the fuzziness the RDD becomes an instrumental variable (IV) estimation, i.e., a two-stage least squares (2SLS) regression (Angrist & Pischke, 2014). Furthermore, the effect of the funds on income inequality will be captured by the income quintile share ratio, which measures the ratio of the total income received by the top 20% of income earners to the total income received by the bottom 20% of income earners (Eurostat, 2021). For the employment analysis, this paper will consider the level of employment within NUTS 2 regions. Lastly, the dataset is constructed using publicly available data provided by the statistical office of the EU.

The findings are inconclusive regarding the relationship between the funding and income inequality within NUTS 2 regions, during the programming periods specified. The magnitude and direction change depending on which bandwidth is applied. Similarly, there does not seem to be any effect of the funding on employment. The results for the employment analysis also vary depending on which bandwidths are used. The findings could either indicate that there does not exist a relationship between the variables or that there is a lack of statistical power due to the small sample size, the latter being more likely for the income inequality analysis. Additionally, there is also a possibility that any effect on employment and income inequality might not materialize within the specific programming periods which could affect the results. Furthermore, by using the income quintile share ratio as the dependent variable the focus is on the top and bottom income quintiles. Other important effects in the middle of the income distributions may be overlooked.

Previous literature has examined the drivers behind the increase in income inequality. While the productivity of workers in Europe has converged over the last decades, the level of unemployment within countries has diverged. This could be one possible explanation for the rise in income inequality (Doran & Jordan, 2013). Furthermore, empirical findings also suggest that openness to trade, technological change and specialization in the workforce are other important drivers. Therefore, high income inequality could be the result of different efforts and talents. Those who work harder and more efficiently earn a higher income leading to greater inequality. On the other hand, the rise in income inequality may also be the result of a less well-functioning economical system, with poorly administered education, inefficient capital markets and the society as a whole being more vulnerable to external shocks (Castells-Quintana et al., 2015).

Earlier studies have also tried to estimate the effects of the cohesion policy and the funds on regional disparities in the EU. However, there does not seem to be a unanimous answer to whether there is an effect in the first place and if that effect is either positive or negative. There is some evidence that the funds do have a positive effect on conditional convergence in terms of GDP/capita (Kyriacou et al., 2014). However, the effect seems to be dependent on the development level of the region. In well-developed regions the effect appears to be negative, while it seems to be positive for less-developed regions (Pinho et al., 2015). In contrast, the well-developed regions still appear to be better at absorbing the funds than their less-developed counterparts (Koudoumakis et al., 2021). Specifically, of the less-developed regions, it appears to be the case that those within well-developed countries experience a larger effect from these funds compared to those within the less-developed countries on the continent (Butkus et al., 2020). However, it is uncertain how large the effect of the funding actually is and a recent study found that while it may be positive for some regions, the effect is surprisingly small. This questions how efficiently the money is spent (Vukašina et al., 2022).

There has also been some discussion whether there is a threshold effect of the funds. This discussion has been centered around the incentives that arise due to the eligibility criteria for receiving funding. As a region must be "sufficiently poor" in order to receive it, there is concern that this will affect the governing politicians to keep the region from developing, as otherwise, there is a risk of them losing funding to their region. This implies that once a region reaches a certain level in development, additional funding might not be used as efficiently to further advance the region. There is evidence that such a threshold effect does exist. In a study between the period 1995 to 2006 it was found that funding of around 1.4% to 1.6% of GDP did reduce disparities, but after that level the disparities were instead aggravated by the funding (Kyriacou & Roca-Sagalés, 2012).

A potential example of this concern materializing is a finding which suggests that the funds may be misallocated. Being a less developed country was found to have a negative correlation with spending on R&D, which is a crucial factor for sustained growth both now and in the future. On the other hand, these countries instead spent large amounts of their funding on physical infrastructure and institution building. However these categories could be argued to be less important for a country trying to achieve sustained economic growth (Medve-Bálint, 2018).

Other factors which also affect the absorption level and in turn the effect of the funds are the pre-existing economic conditions in each region such as institutional quality and educational level. Given the heterogeneity of the regions within the EU, the effect will also be heterogeneous (Kersan-Škabić & Tijanić, 2017; Butkus et al., 2020; Medve-Bálint, 2018). This can explain why the existing studies, that oftentimes use varying samples, reach different conclusions. When mapping the effect in each region across the EU the effect appears to be mostly positive, but there are some regions that have experienced negative effects. Interestingly, there does not necessarily seem to be a connection between the amount of funding and the outcome as there are some regions who receive little funding and yet they still have a significant impact. On the other hand, there are also regions who receive larger amounts of funding with mixed results (Di Caro & Fratesi, 2021).

Thus, it seems as if the cohesion policy does have some effects on disparities in terms of regional GDP/capita, even if these effects may not be homogenous across the EU. However, to the best of our knowledge, all previous literature has mainly focused on regional GDP/capita. Thus, there are no earlier studies that have tried to measure the effect of the funds on regional income inequality. As income inequality plays an important role in the functioning of society, this is an important topic to consider. Given the difficulties of reaching an unanimous conclusion regarding the effect of the funds, it is valuable to examine additional ways in which the effect of the funds can be measured.

The contribution of this paper is twofold. Firstly, it is the utilization of the income quintile share ratio to capture income inequality since it is a novel measurement in the context of the EU and the effect of the funds. In the literature that exists on the funding, most studies have used variants of GDP/capita in order to examine the regional disparities. Secondly, this paper aims at providing external validity to the prior research on the effect of the funds on employment in EU regions.

The paper is structured as follows. In the next section the creation of the EU and the EU's cohesion policy will be described, followed by an overview of how the current cohesion policy is set up. The section will end with an overview of growth and conversion theories. In section three, the boundaries of the study will be outlined, followed by a discussion on what this paper contributes to existing literature and what research questions this paper seeks to answer. In section four a description of the data will be given. In section five the empirical methodology to investigate the research question will be described. In section six the results and a sensitivity analysis will be presented. Section seven will discuss the results in the light of previous literature, convergence theories and the empirical methodology used. The section will also discuss the avenues that exist for future research. The concluding section will summarize the key findings.

2. Background

2.1 History about the cohesion policy

The EU started out as the European Coal and Steel Community, which was established after the end of the second world war. Its primary goal was to create a common market for steel and coal, which had been one of the main industries that created tension between countries. The reasoning was that if this market became more integrated between the countries in Europe, then it would deepen the dependence to such an extent that the risk of a potential future war would be minimized (European Union, n.a.a.). The union became a success and just a few years later it was decided that it would be extended to cover other economic sectors as well, which marked the founding of the European Economic Community, EEC (European Union, n.a.b.). Simultaneously, the European Social Fund (ESF) was also created. The initial aim of this fund was to enhance the social policy of the member countries by enabling its citizens to relocate to new areas within the union and by further strengthening their abilities to take on new occupations. This was seen as an important part of the economic integration that the EEC aimed towards establishing. The Single European Act was accepted in 1987. It aimed to make the member countries' foreign policy more aligned and it also marked the creation of the European Regional Development Fund (ERDF). The aim of this fund was to improve the economic and social cohesion by specifically targeting economically lagging areas (Gabel, 2022).

The budget for the funding has increased significantly since the creation of it. From a total of 64 billion European Currency Units, equivalent to the euro, during the first programming period, the funding has increased to a total of €351 billion for the most recent programming period that ended in 2020. The focus and aim of the funding has shifted slightly between the different programming periods. Growth, employment and innovation has been given an even larger emphasis and recently sustainability and social inclusion has also been incorporated (European Commission, n.a.b.).

2.2 EU's cohesion policy

One of the many goals that the EU has is to increase the cohesion within Europe, i.e., to reduce differences between EU member states. To achieve this, the EU has five funds which are collectively known as the European Structural and Investment Funds (ESIF) (European Commission, 2015a). The funding programme runs over a seven-year period and the budget is set in advance. For each period, depending on current needs, the priorities of the funds may change slightly (Staehr & Urke, 2022). The five funds which constitute the ESIF are the European Regional Development Fund (ERDF), the European Social Fund (ESF), the Cohesion Fund (CF), the European Agricultural Fund for Rural Development (EAFRD), and the European Maritime and Fisheries Fund (EMFF) (European Commission, 2015a).

The ERDF's main aim is to contribute to the development level in less-developed regions in order to reduce imbalances. The fund will do this through promoting innovation and digitalisation among mid-sized companies. It will also support sustainable development, more efficient employment and education, and local development initiatives. The ESF aims toward investing in education and supporting job seekers in Europe through different employment-related projects. The CF specifically targets investments within the transportation network across Europe and investments which have the environment as objective. The EAFRD main aim is toward supporting the countryside, ensuring that they have a sustainable development and that jobs are created and preserved. The EMFF targets the fishing industry and specifically supports individual fishermen and their communities by ensuring a diversified economy and sustainable fishing practices (European Commission, 2015b).

The EU has divided its member states into different statistical regions depending on a set of population thresholds, according to the Nomenclature of Territorial Units for Statistics (NUTS). It is the EU's system for dividing up the economic territory with the purpose of facilitating the collection of regional data, for socio-economic analyses of regions, and to have a systematic sectioning to base the EU regional policies on. There are three NUTS classifications NUTS 1, NUTS 2 and NUTS 3. NUTS 1 are larger socio-economic areas where the population spans between 3 million and 7 million. NUTS 2 regions are the level at which many policies are based upon and the population threshold in these regions are between 800 000 and 3 million. NUTS 3 are even smaller regions with populations ranging from 150 000 to 800 000. The funds are allocated upon a NUTS 2 level (Eurostat, n.a.a.; Eurostat, n.a.c.)

This thesis will focus on two programming periods, namely the period between 2007 to 2013 and the period between 2014 to 2020. During these two programming periods, the regions that were mainly targeted were those with a GDP/capita that was below 75% of the regional average GDP/capita, which received the largest proportion of funds (European Commission, n.a.c.; European Commission, n.a.d.). The funds that targeted these regions were the ERDF and ESF, known as the structural funds. In the last programming period, the EAFRD was also granted based upon the same threshold (European Commission, 2007; European Commission, 2015a; Nordregio, n.a.). This method has been used by previous studies in order to examine the effect of the funds (Becker et al., 2010; Butkus et al., 2019; Mohl & Hagen, 2010).

2.3 Growth and convergence theories

2.3.1 Solow growth model

The Solow growth model was developed by Robert Solow and it tries to explain the factors that lead to economic growth. Through a simple production function, the amount of capital, labor and the level of total factor productivity determines the output in an economy. Capital and labor are endogenously given while total factor productivity is exogenous. However, the most important feature of the Solow model is that capital is assumed to have decreasing returns to scale. Thus, in regions where the amount of capital per worker is high, additional capital will increase the output less than in regions with lower amounts of capital per worker.

Furthermore, the model also exhibits constant returns to scale regarding depreciation of capital. This causes the model to have a steady state which it will always move towards until it reaches that point. Furthermore, if the total factor productivity, e.g. technology, and the population growth rate is the same between two countries, the Solow model assumes that they share the same steady state. Thus, the Solow model assumes conditional convergence. The decreasing returns to scale of capital also implies that once poorer regions increase their investments, their GDP/capita will grow much faster than a similar investment would cause in rich countries. Thus, the poor regions would experience a catching-up effect with the richer regions in such a situation (Corporate Finance Institute, 2022).

2.3.2 Endogenous growth model

The endogenous growth model, or sometimes known as the Romer model takes a different approach to explaining economic growth. According to this model, ideas are non-rivalrous, meaning that they cannot be sold on a competitive market because their marginal cost is very small. In other words, ideas and innovations are the driving forces of economic growth in this model. Due to the accumulation and spillover effects of ideas and innovations, it leads to increasing returns to scale since the idea generation is self-reinforcing. Thus, the endogenous growth model predicts divergence between countries as those who are more innovative will be able to leverage their already existing ideas. Thus, this model may be an explanation for the persistent gaps between countries (Romer, 1990).

2.3.3 New economic geography theory

The new economic geography theory predicts that economic activity will concentrate in specific regions. This concentration happens because proximity reduces transportation and communication costs. Additionally, concentration may also lead to increasing returns to scale due to, for example, spillover effects of knowledge or public goods which companies in near proximity can draw benefits from, such as city centers. Regions with more firms may also have a larger workforce thus enabling the individual companies to find more highly skilled people within these regions. Therefore, if there are increasing returns to scale and low transportation costs, then it may be more profitable to gather the production to one location, which will create a clustering effect (Venables, 2016).

2.3.4 Heckscher-Ohlin trade model

According to the Heckscher-Ohlin trade model, international trade will lead to greater income differences between the highly skilled and the low skilled workers within a developed country. On the other hand, international trade will lead to less income differences within a less developed country as the low income workers become more demanded and thus earn a higher income than before (Castells-Quintana et al., 2015).

Openness to trade has been found to have an association with higher levels of income inequality. However, low-income countries have been observed to be the main losers in Europe, something which contradicts the Heckscher-Ohlin model (Ezcurra, 2019). One possible explanation for this development in Europe could potentially be skill-biased technological change, which both benefits the highly skilled and the low-skilled, but is disadvantageous towards the middle-skilled workers. The highly skilled can increase their productivity due to the new technologies, while the low skilled remain unaffected since they mainly work with human interactions which may be difficult or not profitable enough to replace. The middle-skilled workers however, who perform routine tasks which easily can be automatized may face the risk of losing their jobs. Thus, if the low-income countries in Europe have many middle-skilled workers, the skill-biased technological change may possibly explain why these countries seem to be the main losers (Castells-Quintana et al., 2015).

2.4 The relationship between income inequality and economic growth

Income inequality is interesting as it has an inevitable connection to the functioning of society, as seen above. The studies that exist on the relationship between income inequality and economic growth do not present a unanimous answer. However, one specific pattern appears when examining earlier literature on the topic, which is the development level of the country. In general, it appears as if less developed countries in the world are more adversely affected by higher income inequality levels compared to more developed countries (Gründler & Scheuermeyer, 2018; Neves et al., 2016).

The different transmission channels through which high income inequality negatively affects economic growth appears to be credit constraints, limiting individual's possibilities to education, greater political instability in society and increased fertility rate (Neves et al., 2016). Additionally, the link between income inequality and political instability is especially evident when the institutions are of low quality (Malikov & Alimov, 2022). On the flip side, high income inequality may positively affect economic growth through increased savings and R&D (Neves et al., 2016). Wealthier people are prone to save and invest more than less wealthy people. Thus, a society with high income inequality, and therefore many wealthier people, may have a greater amount of investments and therefore see a higher economic growth rate (Gründler & Scheuermeyer, 2018). It is possible that these different channels can explain why high income inequality appears to have varying effects depending on the development level of the country.

Another way to explain the relationship between income inequality and economic growth is through the Kuznets curve. It was developed by Simon Kuznets during the 1950s. It suggests that once a country starts to develop, income inequality will initially rise until it reaches a certain threshold and then gradually decline as the economic progress continues. Thus, the development pattern resembles an inverted U-curve. The reasoning behind this relationship is that during the initial development, the economic activity is believed to be concentrated in certain areas and then as the development continues, the acquired technology, infrastructure and skills diffuse into the whole region or country (Kuznets, 1955).

Using EU data some studies have found and confirmed the Kuznets curve relationship (Ezcurra, 2019; Soava et al., 2019). The more developed EU countries saw a positive growth effect from inequality while it had a negative effect for the less developed EU countries (Jianu et al., 2021). However, as all countries in the EU could be considered as developed by world standards, it is argued that by using EU-data it should only be possible to observe the latter part of the inverted U-curve, i.e. a downward trend toward less income inequality. This does not align with what is observed in the EU since there is an increasing trend in income inequality within member states. Thus, the inverted U-curve may in fact be an N-shaped curve as developed countries who experience further technological progress may see an increase in inequality (Castells-Quintana et al., 2015).

In sum, while there may not be an exact answer to how income inequality affects economic growth, it appears as if the level of development is vital to examine the impact (Jianu et al., 2021; Shin, 2012; Soava et al., 2019). Income inequality can be a positive indicator as it may be associated with an economy with high levels of investments and R&D. However, as seen above this may only be true for already developed economies, while those who are developing may be better off if the income inequality level is lower.

3. Research specification

3.1 Delimitations

This research paper will focus on a time period of 14 years, from the beginning of 2007 to the end of 2020. During this time period the funds were distributed in two periods: one period from 2007 to 2013 and the other period from 2014 to 2020. Furthermore, as described in the background, the EU has divided its territory into NUTS 1, NUTS 2, and NUTS 3 regions, and the funds are distributed at the NUTS 2 level. Thus, the research will be delimited to the effect within NUTS 2 regions. Furthermore, we will only examine the funds which are administered through the threshold mechanism.

3.2 Academic contribution

This paper will examine if the funds provided by the EU has had an impact on income inequality and employment over the period 2007 to 2020. To the best of our knowledge, there is no earlier literature which has tried to measure any potential effect of the funds on income inequality. A variable that has not been examined in earlier literature is the income quintile share ratio which measures the income inequality. Previous literature has mainly focused on measuring the effect of the funds on GDP/capita. However, income inequality is also an important factor for the cohesion of society and in extension Europe. It is thus beneficial to examine if, and how, income inequality is affected by the funds provided by the EU. This will also serve as an additional way to evaluate the EU policy regarding its funding. This is valuable considering the difficulties of previous research to draw a definite conclusion regarding the effects of the funding. Furthermore, the vast majority of prior studies on the impact of the funds on growth has been concentrated to the time period before 2009 (Butkus et al., 2019; Mohl & Hagen, 2010). This paper will add to previous literature by investigating the more recent years.

Similarly as previous studies assessing the impact of the funding, this paper will use the threshold at 75% in order to assess the effect (Becker et al., 2010; Butkus et al., 2019). Thus, the main question this paper sets out to examine is if the funds have had an impact on income inequality and employment between the years 2007 to 2020. Additionally, considering possible misallocation and threshold effects of funding, this paper will also rule out whether or not the impact of funds are positive or negative. This paper will also add to the validity of prior research on the effect of the funds on employment by analyzing the effect from 2007 to 2020.

3.3 Research questions

1. Has the funds had an impact on the income inequality within NUTS 2 regions in the European Union between the years 2007 to 2020?

2. Has the funds had an impact on the employment levels within NUTS 2 regions in the European Union between the years 2007 to 2020?

4. Data

4.1 Description of the dependent variables

This paper focuses on income inequality and employment. Another central variable is our running variable, which is the initial relative GDP/capita, i.e. the regional GDP/capita relative to the EU average. This variable will be described in the next section.

Income inequality refers to how evenly distributed the income is within a population (IMF, n.a.). There are different ways in which income inequality can be measured. The most prominent of these is the gini coefficient. However, data on the gini coefficient is only available at the country level for the EU and not at a NUTS 2 level. Therefore, we instead use a ratio measure which is available at the NUTS 2 level. The data on income inequality used in this paper is gathered from Eurostat, which is responsible for providing European statistics (Eurostat, n.a.b.). This dataset measures income inequality as a ratio of the top 20% income earners compared to the bottom 20% of income earners. This dataset is called the income quintile share ratio (Eurostat, 2021).

The data for the dependent variable that will capture changes in employment was also gathered from Eurostat. This employment data measures the number of people employed between the ages 15 to 64 on a NUTS 2 level and the measurement is in thousands of people. For measuring the impact of the funding on income inequality and employment, we used the average change of income inequality and employment across the whole programming period.

4.2 Data description

Table 1 shows a summary over the total number of NUTS 2 regions during the two programming periods. Before each programming period, the EU redefines what should be emphasized and focused on, which thus implies that there are some changes between periods. During the first programming period, regions eligible for funding were referred to as objective 1 regions. During the second programming period, these regions are instead referred to as less developed regions (European Commission, 2007; European Commission, 2015b). In this paper, we will refer to these regions as "eligible regions".

Table 1: Eligible regions

	2007-2013	2014-2020
NUTS 2		
Total number of NUTS 2 regions	273	276
Number of eligible NUTS 2 regions	88	72

Sources: (European Commission, 2015b; European Commission, 2007; Eurostat, 2007; Eurostat, 2015)

Data for table 1 regarding the total number of NUTS 2 regions is gathered from the reports on regions in the EU from 2006 and from 2013 (Eurostat, 2007; Eurostat, 2015). Data over the eligible regions is gathered from the EU policy reports (European Commission, 2007; European Commission, 2015b)

The paper written by Becker et al. (2010) mentioned that there may be some discrepancy between regions who received funding and regions who should have received funding according to the 75% GDP/capita threshold. Thus, to examine this possible discrepancy, this paper also looks into this by comparing regions receiving funding according to the EU and our own calculations over which regions should have received funding according to the 75% threshold (European Commission 2007; European Commission, 2015b).

As there is no publicly available data on GDP/capita per NUTS 2 region to the extent that was required, this was calculated by us. Data over the population and GDP per NUTS 2 region was collected from Eurostat (Eurostat, 2022; Eurostat, 2023). These two datasets were then combined in order to calculate the GDP/capita per region. Through these calculations, an average GDP/capita per region was calculated for the base years 2006 and 2013. Similarly to Becker et al. (2010), the year before the start of the programming period was used in order to evaluate whether a country should have received funding or not. This was then compared to the list of regions receiving funding according to the European commission (European Commission 2007; European Commission, 2015b).

Figure 2 and 3 in the appendix shows a visual presentation over which regions received funding. Green areas represent regions which had a GDP/capita below 75%. The yellow areas are regions with a GDP/capita between 75-80%. The orange areas are regions with GDP/capita between 80-85% and the red areas are regions that had a GDP/capita above 85%. There is some discrepancy seen in figure 2 and 3 as the EU also had some "phasing-out" regions that were given transitional support even though their GDP/capita was between 75-90% of the average GDP/capita of the EU (European Commission, 2007; European Commission, 2015b). In order to test how this affects the results, we performed two separate analyses. In the first analysis, which is presented under the result section, we include the regions given transitional support as treated regions. In the second analysis, which is presented under the sensitivity checks, we view the regions given transitional support as untreated. Earlier literature does not mention these transition regions and therefore, we wanted to test how it would affect the result if these were viewed as untreated.

Table 2a: Eligibility and treatment according to the 75% GDP/capita threshold

	Recipients	Non-recipients
	NUTS 2	NUTS 2
2007-2013 EU		
Eligible	72	3
Non eligible	16*	184

Notes: The former NUTS 2 regions Dessau and Magdeburg merged into a larger NUTS 2 region. Thus, there is no available data to calculate if these two regions were above or below the threshold. Therefore, we assumed that they followed a similar pattern as the other German NUTS 2, namely that they were above the threshold.
Sources: (European Commission, 2015b; European Commission, 2007; Eurostat, 2023a; Eurostat, 2022)

Table 2b: Eligibility and treatment according to the 75% GDP/capita threshold

	Recipients	Non-recipients
	NUTS 2	NUTS 2
2014-2020 EU		
Eligible	66	7
Non eligible	6	197

Sources: (European Commission, 2015b; European Commission, 2007; Eurostat, 2007; Eurostat, 2015)

Table 2a and table 2b illustrate a summary regarding the discrepancy discussed earlier. We constructed it by comparing the recipients of funding with our own calculations regarding which regions should have been eligible and not according to the 75% GDP/capita threshold. Table 3 and 4 summarizes the average, minimum and maximum income inequality and employment levels for the regions within each country. Table 5, 6 and 7 in the appendix shows an extensive list for all the regions within the different countries which we have used. Due to missing data on income inequality and on employment, some regions were filtered out. This also implies that there is a risk of hidden biases in the data sample. In total, we used three datasets to test whether the funding had a significant impact. The first dataset, used for the income inequality analysis, consists of the regions listed in table 5. To test the effect on employment, we matched the data to the regions used for income inequality. However, a few regions are missing, see more in table 6. Since more data was accessible for the employment level, we also created a larger dataset using all regions available. See more about this in table 7.

5. Empirical Methodology: Fuzzy Regression Discontinuity Design

There are situations in the real world which lend themselves to being useful in experimental designs. These situations can arise when seemingly arbitrary rules cause a cutoff point to occur. Due to the rule being arbitrary and not connected to something else, the only thing that should change at the cutoff is the rule. In our case, the rule is the 75% threshold which entitles regions within the EU to support from the funds. Near the threshold, i.e. between the regions just below and the ones just above, the only thing that reasonably should differ is the access to extra funding through this scheme. If there is a significant difference in income inequality or employment in combination with no other changes near the threshold, then it may be reasonable to draw the conclusion that access to funding impacts the employment or income inequality. This is what is known as an RDD (Angrist & Pischke, 2014).

Thus, in an RDD study, the threshold is vital in order to study the effect. The variable which determines when the cutoff is crossed is called the running variable. In our case, the running variable is the relative GDP/capita. An RDD study requires that it is solely the rule which determines whether or not an entity is being treated. Furthermore, up until the threshold is crossed, there is no treatment effect taking place no matter how near the entity is to the cutoff point. This is what is known as a sharp regression discontinuity (RD). However, there are cases when the threshold significantly impacts the likelihood of treatment but not to the full extent as in a sharp RD and this is instead known as a fuzzy RD. In our case, there are eligible regions who did not get funding and non-eligible regions which did get funding. This implies that we have to use a fuzzy RDD (Angrist & Pischke, 2014).

The coefficient that would measure the effect of the threshold in a sharp RDD instead measures the intent-to-treat effect in a fuzzy RD. The reduced form regression seen below is the regression that we want to estimate, which is the effect stemming from the higher probability of treatment. This is typically done through a 2SLS regression (Angrist & Pischke, 2014). We will estimate how the income inequality and employment is affected by access to the funds. Thus, our reduced form regressions measures how income inequality and employment are affected by the cutoff variable, which is a dummy indicating whether GDP/capita is below or above the 75% threshold and the running variable GDP which represents the GDP/capita in relation to the average GDP/capita.

Reduced form: $Income\ Inequality_{it} = \alpha_0 + \beta_1 Cutoff + \beta_2 GDP + \epsilon_{it}$ (1)

Reduced form: $Employment_{it} = \alpha_0 + \beta_1 Cutoff + \beta_2 GDP + \epsilon_{it}$ (2)

The cutoff variable is defined as seen below:

$Cutoff_{it} = 1\ if\ GDP \leq 75\%$

$Cutoff_{it} = 0\ if\ GDP > 75\%$

Thus, in our 2SLS regression, the first stage will estimate the treatment variable based on the cutoff dummy and the running variable as seen below.

First Stage: $\widehat{Treatment} = \gamma_0 + \gamma_1 Cutoff_{it} + \gamma_2 GDP_{it} + \omega_{it}$ (3)

The estimated treatment variable will then be used in our second stage where we can measure the effect of the treatment on income inequality and employment. Therefore, a fuzzy RDD uses an IV approach where the treatment variable is estimated in the first regression using an instrument, which in this case is the cutoff. The treatment variable is then used in the second regression to estimate the effect on the dependent variable (Stewart, 2016).

Second Stage: $Income\ Inequality_{it} = \beta_0 + \beta_1 \widehat{Treatment}_{it} + \beta_2 GDP_i + \epsilon_{it}$ (4)

Second Stage: $Employment_{it} = \beta_0 + \beta_1 \widehat{Treatment}_{it} + \beta_2 GDP_i + \epsilon_{it}$ (5)

By examining the coefficient from the treatment variable in the second stage regression, we will see whether there was an effect on income inequality and employment. We will also use different bandwidths and examine if there are significant changes near the threshold in other variables which may be connected. Since the fuzzy RDD is the same as an IV regression, there are three requirements which must hold. These three are:

1. Instrument relevance: $corr(Z_i, X_i) \neq 0$
2. Instrument exogeneity: $corr(Z_i, u_i) = 0$
3. Instrument exclusion: $corr(Z_i, X_i \mid X_i) = 0$

The first assumption implies that the cutoff creates access to the treatment. In our case, the cutoff is the 75% threshold and it implies that a region is eligible for treatment, i.e. the funding. The second assumption implies that there should be no correlation between the cutoff threshold and other factors which may influence the dependent variable. Given that the threshold is rather arbitrary, other factors which also influence income inequality and employment levels should not be correlated to it (Stock & Watson, 2019). The last assumption implies that the cutoff affects the employment level and income inequality only through the access to the treatment. While the regional GDP/capita may be correlated to the average income inequality and employment levels, the specific cutoff itself should not be correlated as it is a rather arbitrary threshold (Heiss, 2020).

We follow the article *Going NUTS: The effect of EU Structural Funds on Regional Performance* written by Becker et al. (2010) who used a similar method in order to evaluate the impact of the funds. The specific threshold suggests that it is possible to draw benefit from this using an RDD study (Stock & Watson, 2019). The design is built upon the idea that near the threshold, the entities should not differ substantially and thus an RDD should produce unbiased estimates without the need for control variables (Holla, n.a.). This is an advantage as there may be many other factors which can affect the dependent variable but through an RDD, this possible error is eliminated. Therefore, it should be easier to claim causality.

However, this also implies that while the internal validity of an RDD study may be strong, the external validity may be rather weak. The results from an sharp RDD study are mainly valid for the subpopulation near the threshold and thus they do not tell much about the average treatment effect (ATE), even if the sample would be large. In a fuzzy RDD, the IV approach implies that it is the conditional average treatment effect for compliers (CATEC) which is measured. This tells us about the effect on the compilers, i.e. the entities who follow the rule, but not the effect on the non-compliers. The result given by the fuzzy RDD for this subpopulation may not have external validity and any generalizations made should be interpreted carefully (Wing & Bello-Gomez, 2018).

The RDD has previously been used in similar studies that exploits the threshold set up by the EU. It is upon this foundation that we have decided to use this method. Other methods such as the GMM estimator, spatio-temporal econometric model, DID estimator, OLS, and the FE model have also been used to study the effects of the funding (Butkus et al., 2019; Mohl & Hagen, 2010). We did consider an OLS model with control variables and a FE model. However, an OLS model was deemed to be suboptimal due the difficulties in determining which control variables to include. Given the many factors which may influence income inequality and employment, having the right controls would have been crucial in order to deal with omitted variable bias. The FE model would have controlled for entity and time fixed effects (Stock & Watson, 2019). However, as the funding is set up in a specific way, this is suitable for the fuzzy RDD while an FE model would not exploit the nature of this threshold. Thus, we deemed it appropriate to apply an RDD to evaluate the effects of the funds on employment and income inequality.

6. Results

6.1 Checking for fuzziness

In order to establish that the regression discontinuity is fuzzy, the compliance was plotted in order to visually observe whether our dataset did consist of non-compliers, see graph 1 in the appendix.

Table 8: Eligibility and treatment according to the 75% GDP/ capita threshold

	Recipients	Non-recipients
Income inequality		
	NUTS 2	NUTS 2
Eligible	60	2
Non eligible	14	70
Employment - matched		
Eligible	58	2
Non eligible	12	70
Employment - full		
Eligible	139	8
Non eligible	69	284

Sources: (European Commission, 2015b; European Commission, 2007)

Table 8 above shows the compliance for the three datasets. All the datasets show a similar pattern where the rules for being eligible and receiving funding were violated for some regions. Furthermore, it can be observed that the amount of non eligible regions which received funding were higher than the number of eligible regions that did not receive funding.

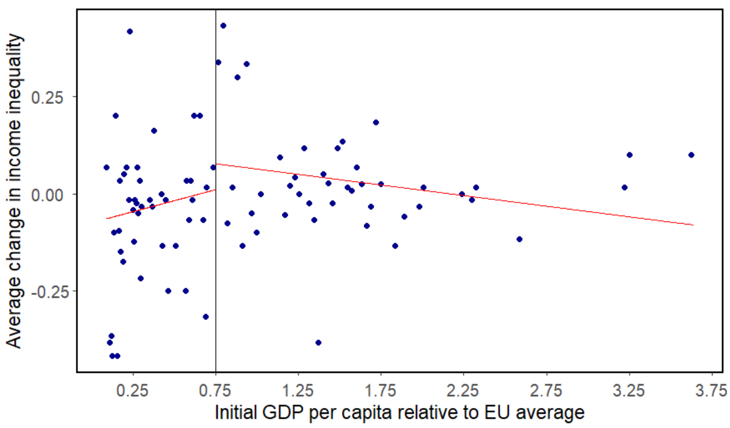
6.2 Regression discontinuity results

For each dataset there are four regressions, each with different bandwidths. The first regression incorporates the entire dataset and the rest of the regressions has the following bandwidths: +/- 70%, +/- 50% and +/- 20%. The bandwidth limits the amount of observations around the threshold of 75% by excluding the values outside of it.

6.2.1 Income inequality

This section presents the result for the fuzzy regression discontinuity on the average change in income inequality. Firstly, one can visually observe in graph 2 below that there exists a slight discontinuity at the threshold in the full sample. This would suggest that the regions just above the threshold, on average, have had a slightly higher level of change in income inequality than the regions just below the threshold. Secondly, it appears as if the slight upwards trend is changed to a slight downward trend after passing the threshold. This change in slope would suggest that income inequality increases until the threshold is reached and then decreases after that point. However, as can be seen in the graph below, there are some outliers which may affect the results. Thus, in order to test the robustness of these initial results, we restricted the data sample in graph 3, which can be seen in the appendix. The linear trends seen below change when restricting the sample, indicating that the outliers may affect the results. In the restricted sample, there is no trend before the threshold and a very modest upward trend after it. Thus, the funding may not have any significant impact on income inequality.

Graph 2: Graphical results for income inequality



Sources: (Eurostat, 2021; Eurostat, 2022; Eurostat, 2023a; European Commission, 2015b; European Commission, 2007)

The results of the regression in table 9 below shows that the treatment coefficient is negative in all except the last regression with the bandwidth of +/- 20%. A negative coefficient would suggest that the treatment has decreased the income inequality. The regression with a bandwidth of +/- 20% yields the opposite result. The average level of income inequality in our sample is 4.54 and thus a change of 1.24 in income inequality would be equal to a 27.3% change. However, the p-value indicates that the results are not significant in any of the four regressions. Additionally, the large standard deviations in combination with the change of sign of the coefficient further shows that these results are not statistically significant.

Table 9: Regression discontinuity results on income inequality

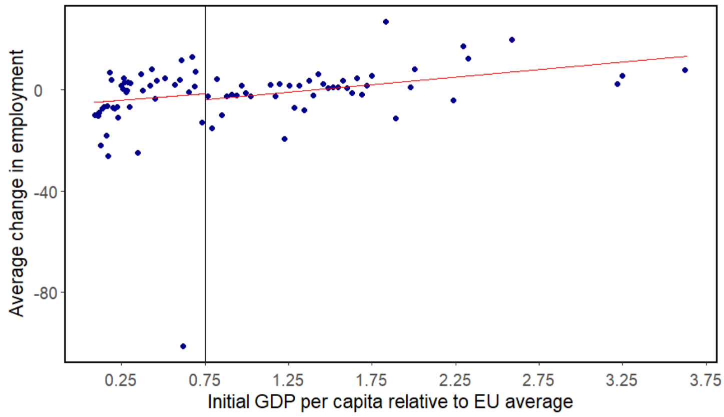
Bandwidths	None	+/- 70%	+/- 50%	+/- 20%
Intercept	0.16	1.43	1.91	-0.84
	(0.12)	(3.69)	(5.38)	(1.05)
GDP	-0.11	-1.85	-2.24	2.12
	(0.10)	(4.91)	(6.59)	(1.80)
Treatment	-0.25	-2.29	-2.77	1.24
	(0.17)	(5.90)	(7.86)	(1.34)
Observations	146	107	70	26
p-value *of Treatment*	0.15	0.70	0.73	0.36

Notes: *** p < 0.001; ** p < 0.01; * p < 0.05
Sources: (Eurostat, 2021; Eurostat, 2022; Eurostat, 2023a; European Commission, 2015b; European Commission 2007)

6.2.2 Employment: matched

This section presents the result for the fuzzy regression discontinuity on the average change in employment when using the matched employment dataset. In graph 4 below, it seems as if there is no discontinuity at the threshold and no change in the trend either before or after the threshold. In order to follow the same methodology as above, we restricted the sample in order to see if these results were robust. This can be seen in graph 3 in the appendix. When restricting the sample, one can see that the trend before the threshold is steeper than after. In both cases the slope is positive. Thus, prior to the threshold, the regions experience a more significant change in employment level compared to after the threshold.

Graph 4: Graphical results for matched employment analysis



Sources: (Eurostat, 2022; Eurostat, 2023a; Eurostat, 2023b; European Commission, 2015b; European Commission, 2007)

The results for the regressions on the matched employment dataset are presented in table 10 below and it shows that the treatment coefficient is positive in all regressions except the last one. A positive coefficient could imply that the treatment increases the level of employment, while a negative coefficient implies that treatment decreases it.

Table 10: Regression discontinuity results on employment matched

Bandwidths	None	+/- 70%	+/- 50%	+/- 20%
Intercept	-7.03	-33.40	-8.54	35.62
	(4.98)	(72.2)	(62.6)	(47.08)
GDP	8.64	45.27	3.68	-69.2
	(4.79)	(100.3)	(82.61)	(116.6)
Treatment	7.46	50.43	9.28	-57.23
	(8.57)	(119.38)	(96.3)	(63.0)
Observations	142	103	66	22
p-value *of treatment*	0.39	0.67	0.92	0.37

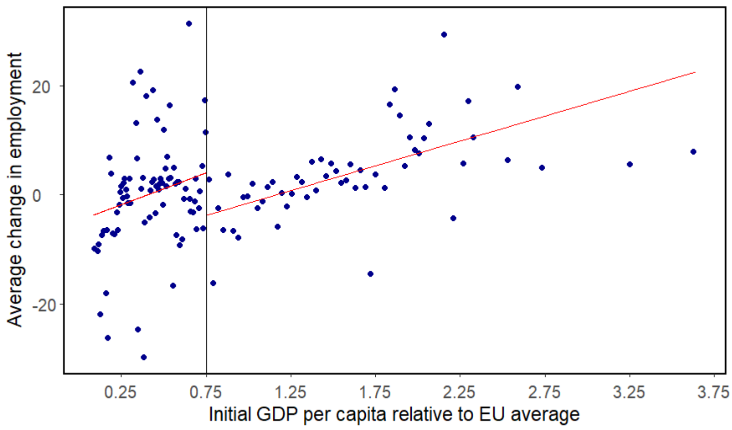
Notes: *** p < 0.001; ** p < 0.01; * p < 0.05
Sources: (Eurostat, 2022; Eurostat, 2023a; Eurostat, 2023b; European Commission, 2015b; European Commission, 2007)

However, the p-value of the table above indicates that these results are not significant and there is a great variation in the magnitude depending on which bandwidth is used. Additionally, the large standard deviations in combination with the change in sign of the coefficient further shows that these results are not statistically significant.

6.2.3 Employment: full

This section presents the result for the fuzzy regression discontinuity on the average change in employment when using the dataset with all available observations. Through visually analyzing graph 5 below it can be noted that the trend before the threshold is positive. At the threshold the line jumps slightly downwards and the slope remains similar. This would suggest that the treatment does increase the employment level in NUTS 2 regions. In order to test the robustness of this, we restricted the sample in the same manner as above. This can be seen in graph 3 in the appendix. When restricting the sample, the same pattern emerges which would support the robustness of the results. However, we also need to take into account the regression results when using different bandwidths.

Graph 5: Graphical results for full employment analysis



Sources: (Eurostat, 2022; Eurostat, 2023a; Eurostat, 2023b; European Commission, 2015b; European Commission, 2007)

Table 11 shows that the treatment coefficient is positive for the first two regressions and then changes sign for the last two regressions. A positive coefficient would suggest that treatment increases the employment level, while a negative coefficient suggests that treatment decreases it. The first regression has a p-value at the significance level of 0.01. However, this result is still not significant in practice as no bandwidth is used. The more interesting regressions are those with bandwidths and especially the last one with the strictest one. The p-value of those regressions indicate that there is no statistical significance of the results.

Table 11: Regression discontinuity results on employment full

Bandwidths	None	+/- 70%	+/- 50%	+/- 20%
Intercept	-11.61*	-162.31	168.06	209.7
	(4.63)	(408.55)	(581.55)	(1145.9)
GDP	17.3**	272.1	-298.19	-370.11
	(5.42)	(677.81)	(1024.06)	(1914.6)
Treatment	18.46**	265.48	-265.06	-278.09
	(7.1)	(667.72)	(918.2)	(1501.121)
Observations	500	414	325	109
p-value *of treatment*	0.01	0.69	0.77	0.85

Notes: *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$
Sources: (Eurostat, 2022; Eurostat, 2023a; Eurostat, 2023b; European Commission, 2015b; European Commission, 2007)

6.3 Sensitivity checks

6.2.1 Manipulation testing

The treatment rule depends on a threshold which is publicly known. This implies that there is a risk of manipulation as government officials may have incentives to keep the regional GDP/capita below the threshold. Possible manipulation has previously been mentioned in the literature and as discussed in the introduction, there is also some evidence for it. Therefore, in order to check for manipulation in our datasets, we conduct a specification testing in order to see if there is evidence of this.

Firstly, however, we made a visual test by plotting the distribution of regions with different GDP/capita. These can be seen in graph 6 in the appendix. From these diagrams it is not obvious that there would be any grouping either before or after the threshold. Thus, there should not be any systematic attempts to stay below the threshold in our samples. However, to statistically check that there is no manipulation in the dataset, we also made a density test. The visualizations of the density test can be seen in graph 7 the appendix. The density test estimates whether there is any discontinuity in the running variable. It is a variant of the local linear density estimator and it creates a finely-gridded histogram. Then, using local linear regression, the histogram is smoothened out on each side of the threshold. This is then tested to see if there is a discontinuity around the threshold. The null hypothesis of this density test is that there is no discontinuity (McCrary, 2006). We did a density test estimation for both the sample used for the income inequality analysis and for the extended sample used for the employment analysis. The results of this density estimation can be seen in table 12 below. In both estimates, the p-value is above 0.05 and therefore we cannot reject the null hypothesis. This implies that there should not be any discontinuity in the running variable and therefore we draw the conclusion that there is not any manipulation in our data sample.

Table 12: Density test estimation

Income inequality		
Method	T	P > | T |
Robust	0.2127	0.8315
Threshold = 75%	Left of threshold	Right of threshold
Number of observations	61	85
Bandwidth estimate	0.402	0.402
Employment		
Method	T	P > | T |
Robust	1.1129	0.2658
Threshold = 75%	Left of threshold	Right of threshold
Number of observations	147	353
Bandwidth estimate	0.197	0.27

Sources: (Eurostat, 2021; Eurostat, 2022; Eurostat, 2023a; Eurostat, 2023b; European Commission, 2015b; European Commission, 2007)

6.2.2 RD Plots

In order to check if there are any other factors which may affect our dependent variables, we created RD graphs 8 and 10. Two sets of RD graphs were conducted where the first set is using regions matched to the income inequality analysis data. The second set is using regions matched to the employment analysis data, see more exact information about which regions in table 17 in the appendix. The reasoning behind using two sets of RD graphs is because there may be different patterns in the data depending on which regions are examined. We used the following variables to do the RD graphs: student in short-cycle tertiary education, employment rate, employment in technology and knowledge intensive sectors, population density, population level and R&D personnel. Students in short-cycle tertiary education may join the working force rather soon, compared to other education programmes that may take longer. Therefore, we believed it was appropriate to include this variable as the funding may affect it, which in its turn could possibly be linked to both income inequality and employment. The employment rate variable was dropped from the second RD set as it would measure the same thing as our main results. However, it was included in the first RD set as the employment rate may affect income inequality levels. Employment and knowledge intensive sectors were included in both RD sets as it may affect both employment and income inequality given that jobs in those sectors may be more well-paying. The same reasoning is behind why R&D personnel were included. Population density measures the ratio between the area of the NUTS 2 region and the average annual population while population level measures the annual population number. These two variables were included because they may affect income inequality and employment given that the population density and level may be larger in cities compared to the countryside.

In graph 8, there does seem to be some changes before and after the threshold in most variables. However, upon closer examination most shifts seem to be due to outliers, which is present in all diagrams below. In order to see how the outliers affected the RD plots, we restricted the data of this first set, which can be seen in graph 9 in the appendix. When comparing the restricted RD plots, the linear trends are indeed shifted. Employment in technology and knowledge intensive factors is the only RD plot which did not change substantially.

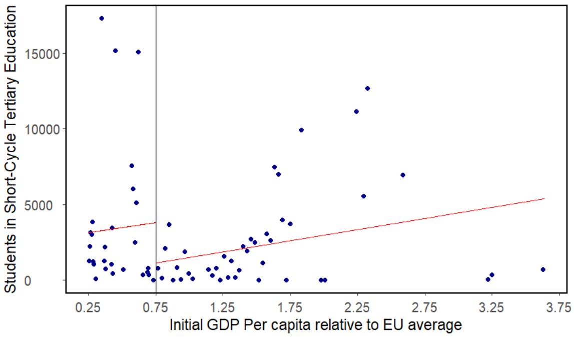
In the second set, we had more observations. There does not seem to be any significant shifts before and after the threshold in any of the variables. In population density there is a slight shift, however it is likely due to the outliers.

In conclusion, there were shifts in the linear trends at the threshold in the sample used for the income inequality analysis. However, there were no such shifts in the sample used for the full employment analysis. One likely explanation could be that there were fewer regions used in the sample for the income inequality analysis which distorts the results. This is also something which may affect the main results. The lack of statistical power is thus something which can be seen in both the RD plots and the main results. The consequence of this is that one cannot draw too strong conclusions from the RD plots and how they affect the main results, nor the main results themselves.

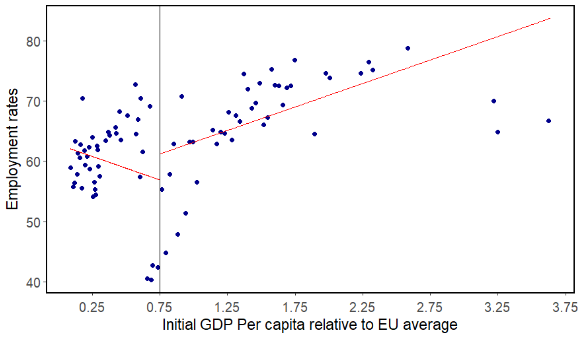
6.2.3 Phasing out regions

The EU has phasing-out regions that were eligible for funding during previous programming periods. These regions have surpassed the threshold but are still given funding through a phasing out scheme. However, earlier literature does not mention these phasing out regions and therefore we wanted to test how the treatment of these regions affect the results. Thus, in the main results we included the phasing out regions as treated regions. Nonetheless, given that earlier literature does not mention these phasing out regions, we wanted to examine how the results are affected if these regions are neglected and viewed as not treated. In table 13, it is possible to see that there are fewer regions classified as recipients and more regions classified as non-recipients with this classification.

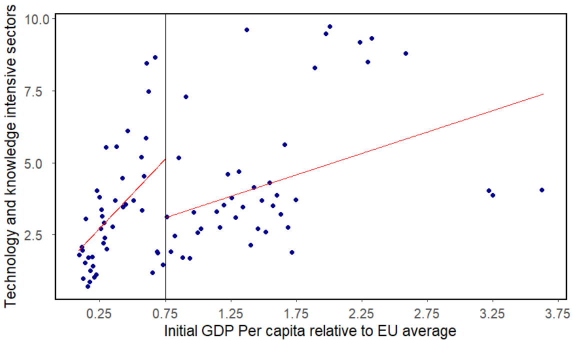
Graph 8: Using regions matched to the income inequality analysis



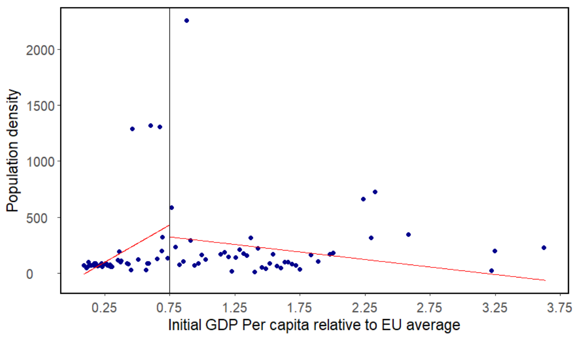
Students in Short-Cycle Tertiary Education



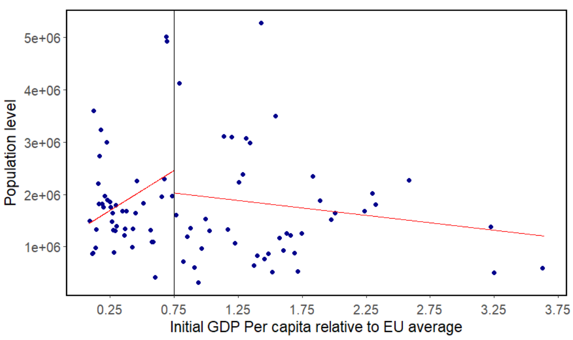
Employment rate



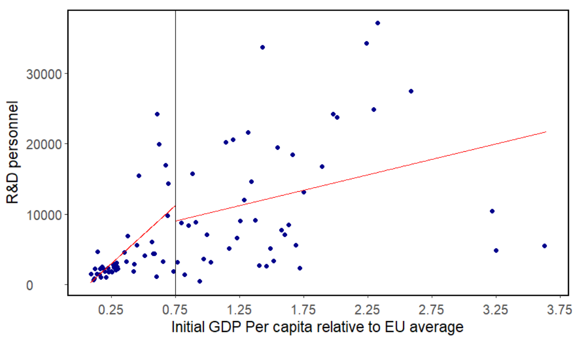
Employment in technology and knowledge intensive sectors



Population density



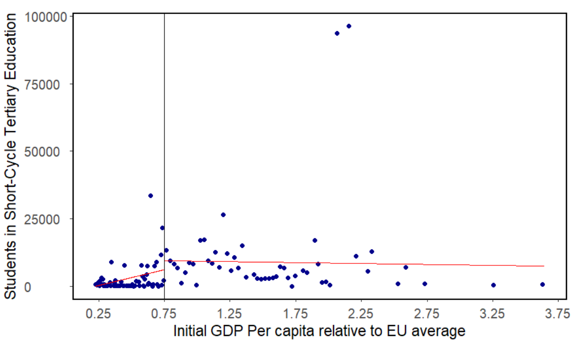
Population level



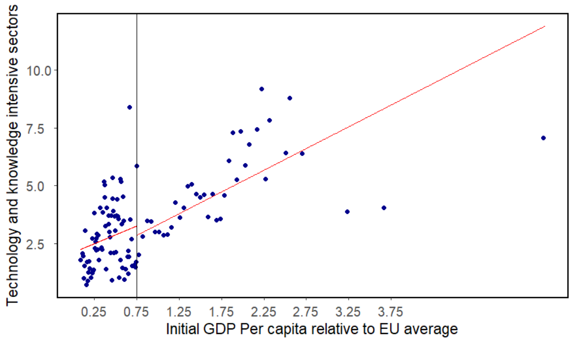
R&D personnel

Sources: (Eurostat, 2022; Eurostat, 2023a; Eurostat, 2023b; Eurostat, 2023c; Eurostat, 2023d; Eurostat, 2023e; Eurostat, 2023f)

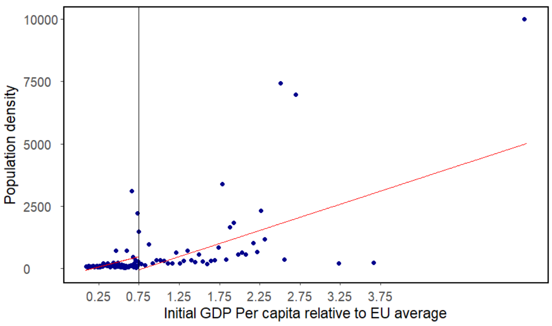
Graph 10: Using regions matched to the employment analysis



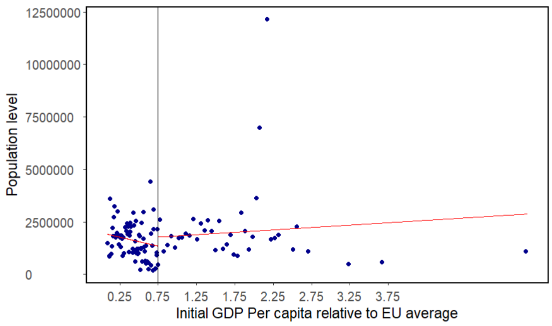
Students in Short-Cycle Tertiary Education



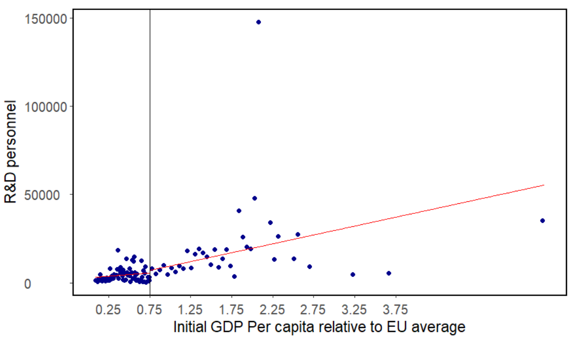
Employment in technology and knowledge intensive sectors



Population density



Population level



R&D personnel

Sources: (Eurostat, 2022; Eurostat, 2023a; Eurostat, 2023b; Eurostat, 2023c; Eurostat, 2023d; Eurostat, 2023e; Eurostat, 2023f)

Table 13: Eligibility and treatment according to the 75% GDP/capita threshold

	Recipients	Non-recipients
Income inequality		
	NUTS 2	NUTS 2
Eligible	58	4
Non eligible	5	79
Employment - matched		
Eligible	56	4
Non eligible	4	78
Employment - full		
Eligible	138	9
Non eligible	57	296

Sources: (Eurostat, 2007; Eurostat, 2015; European Commission, 2015b; European Commission, 2007)

In table 14, 15 and 16 in the appendix, the result for the income inequality and employment analyses are presented where phasing-out regions are neglected. There were some differences when comparing with the main results. In general, the magnitude of the treatment coefficient was lower. The only exception was in the matched employment analysis where the treatment coefficient was larger in magnitude compared to the main results. There were also some differences in the sign of the treatment coefficient since two regressions of the full employment analysis had positive coefficients in contrast to the main results. In general, neglecting the phasing out regions gave lower p-values, however with some exceptions. On the other hand, it was not low enough to make the results significant.

We examine income inequality and employment whilst other research has looked at regional GDP. Thus, we cannot be sure that the treatment of the phasing out regions affects the results in the same manner. It could be the case that the handling of these phasing out regions affect the regional GDP in other ways than it affects income inequality and employment. Furthermore, it may also depend upon the sample used. On the other hand, in our findings, the effect of the funding is not changed by how these phasing out regions are handled. This supports previous research which has found limited effect, at least concerning employment.

7. Discussion

In this paper the effect of the EU funding on income inequality and employment was investigated using a fuzzy RDD approach. Four regressions with different bandwidths for each dataset were used. The results were statistically insignificant, which suggest that no concrete conclusion can be drawn solely based on this study.

There can be different explanations as to why the results are statistically insignificant. The sample size might not have been large enough for the income inequality analysis. The regression with the smallest bandwidth, which is the most interesting to examine, had very few observations. Thus, it could be argued that our income inequality analysis lacked statistical power. It is important to have a lot of data around the threshold in a fuzzy RDD. A larger sample size would probably also have increased the amount of regions near the threshold. Thus, it is likely that the insignificant result for the income inequality might be attributed to the small sample size. However, when analyzing NUTS 2 regions it has to be noted that there are only a limited number of them. Thus, even if the data for the income quintile share ratio would have been complete, the sample size might not have been dramatically different.

The employment analysis with all available data had a larger sample size than the income inequality analysis. Still, there was no statistically significant relationship between the funds and employment. This result is congruent with the findings of Becker et al. (2010). This seems to imply that there is no effect of the funding on employment, however, it could potentially be the case that it takes longer time than one programming period to see an effect. Our RDD is not useful for capturing long-term effects since we only measure the change within one programming period. To examine if any effect occurs over a longer time period, another method would be needed. Nonetheless, the insignificant result raises the question of the funding's efficiency. Additionally, the funding's aim changes slightly over time and between different programming periods. This is something which may affect the result and thus make the measurement of the efficiency of the funding scheme more difficult. Furthermore, this also affects comparability of different studies. Still, we believe that it is important to try and quantify the effect since we deem it important that the money of European taxpayers is spent wisely on policies that drive change.

We do recognize the complexity of measuring the impact of the funding and are aware of the limitations in this paper. The income quintile share ratio focuses on the top and bottom quantiles while it disregards the middle. As discussed in the theory section which describes skill-biased technological change it may be the middle-income workers who lose out the most. By using the income quintile share ratio we miss the effect on this part of the distribution. Thus, it could have been interesting to analyze the middle quantiles as well. Therefore, this is a potential avenue for future research. Additionally, the income inequality dataset and the matched employment dataset used regions from 16 out of the 27 EU member states. This could be a potential source of biases and thus affect the result. However, even though the sample does not include all EU member states, it is worth noting that it does include countries from different geographical areas and with different economical standards. Therefore, the main issue with the data might rather be that the sample is too small.

Previous literature has mentioned that there is an increase of income inequality within countries and an decrease between countries. While our study does not examine the general evolution of income inequality over time, we did not find evidence that income inequality would have been reduced. In accordance with the new economic geography theory, one could possibly argue that countries face a dilemma when given the funding. On the one hand, it may be more efficient to promote growth in the cities given the higher level of skills typically present. On the other hand, promoting growth in the countryside may be less efficient as that typically involves lower-skilled labor. However, at the same time, it would encourage more inclusive growth which could decrease the inequalities within countries. According to this reasoning, it is possible that the funds are spent in ways which may be seen as more efficient, but inadvertently may also spur income inequality through this clustering effect. However, our findings suggest that there is no significant relationship between the funding and income inequality, which can be interpreted as that while the funding does not decrease income inequality, it at least does not aggravate the situation.

The endogenous growth model suggests that ideas are the leading factor behind continuous growth. Furthermore, idea generation creates a positive feedback cycle, which implies that countries which have had a head start will keep their relative position. This explains why there is a persistent gap between countries in regards to development level. In this light, the single market created by the EU is one way in order to facilitate the transmission of ideas across borders. This may have played a part in decreasing the income inequality differences between countries. On the other hand, the new geographic theory may explain why the income inequality differences within countries have increased.

Lastly, the Heckscher-Ohlin theory of international trade predicts that the wealthy will gain in rich countries and lose in the poor countries. On the other hand, the less wealthy will lose in rich countries and gain in the poor countries. This should imply that the less wealthy individuals in the more poor countries in Europe gain from the common market which the EU has established. Income inequality should thus decrease in the less wealthy countries as a consequence of this. However, this model does not take the concept of “brain drain” into account. The wealthy in poor countries can easily move within the EU, which in turn may affect the country’s long term possibilities to achieve economic growth. As previously mentioned in the paper, one of the EU’s main goals is to increase cohesion. Thus, potential “brain drain” may be destructive for the cohesion. In this light, the funding can be seen as a redistributive measure to combat the negative effects of this. However, as this paper finds, there seems to exist no significant relationship between funding and reduced income inequality within receiving regions.

Previous studies have not mentioned the phasing-out regions. Our result indicates that it does not seem to matter how the phasing out regions were treated in the analysis. However, the handling of these regions did affect the p-value slightly and thus there is a risk that they could still bias earlier research. Furthermore, regarding the extrapolation potential of our findings, the internal validity is rather strong for an RDD as discussed in the method section. On the other hand, the external validity of RDD studies is rather weak which is problematic for this purpose. This implies that the result of our findings may not be extrapolatable to other regions within the EU not included in the dataset even though the dataset could be argued to be representative for the whole EU, as well as for similar funding schemes elsewhere in the world. However, there are methods which can be applied in order to increase the external validity. While they are outside the scope of this paper, we do suggest further research to apply such methods in order to increase the external validity of our findings. One such method is comparative RDD which uses a comparison group that did not have any treatment, both above and below the threshold. Thus, in our setting it would involve using data from outside the EU. Other methods which may also increase the external validity are covariate matching RDD, treatment effect derivatives and statistical checks for selection bias (Wing & Bello-Gomez, 2018).

Another limitation of this study, as described above, is the small amount of data near the threshold for the income inequality analysis. Our motivation behind using the RDD was to exploit the specific threshold already existing. Thus, initially we were not aware of the exact amount of data points that would end up close to the threshold. After having conducted the study, we see the potential of using other methods to be able to use all the data and thus make stronger conclusions. Therefore, we also suggest that future research uses other statistical models to broaden the scope of analysis, especially for the income inequality data. One such statistical method is the FE model which would also measure the intensity of funding, which would be able to reveal more about the necessary amounts of funding required to observe a change.

8. Conclusion

This research paper sought to answer whether EU funding has had an impact on the income inequality and employment level within NUTS 2 regions between the years 2007 to 2020. Given the heavy focus on economic growth and the inconclusive results of previous studies, we believe that it is of large interest to increase the scope of analysis regarding the impact of the funding. A fuzzy RDD has been used to investigate the effect of the policy, as previous studies have done. Four regressions with different bandwidths were used in addition to a sensitivity analysis in order to see how robust the results were.

The results of this study suggest that there is no effect on employment levels. Given the larger sample size for the employment analysis, we can conclude that the policy has most likely not been effective in the examined years which is in line with the findings of previous literature. Similarly, the results regarding income inequality were statistically insignificant. However, the sample size was not as large and therefore we can not draw an equally strong conclusion regarding the effectiveness of the funding. Considering the lower sample size for the income inequality analysis and the importance of a high number of observations close to the threshold, we suggest that further research is done using other measures that may be less sensitive to this. The results of this study could be seen as questioning how efficiently the EU funding is spent. Nonetheless, as cohesion is a broad goal, it may be difficult to quantify the effect given the many aspects which it entails. Therefore, we recommend that future research adopt additional measures and investigate other methods as well in order to further examine this highly relevant topic.

Reference list

Alesina, Alberto, Rafael Di Tella, & Robert MacCull. (2014). Inequality and happiness: Are Europeans and Americans different? *Journal of Public Economics, 88*(9-10) doi:10.1016/j.jpubeco.2003.07.006

Angrist, D. J., & Pischke, J. (2014). *Mastering metrics: The path from cause to effect.* Princeton University Press.

Barnier, M. (2003). The future of European regional policy. Retrieved from https://www.intereconomics.eu/pdf-download/year/2003/number/6/article/eu-cohesion-policy-challenges-and-responses.html

Becker, S. O., Egger, P. H., & von Ehrlich, M. (2010). Going NUTS: The effect of EU structural funds on regional performance. *Journal of Public Economics, 94*(9), 578-590. doi: 10.1016/j.jpubeco.2010.06.006

Beckfield, J. (2006). *European Integration and Income Inequality*. American Sociological Review, 71(6), 964–985. doi:10.1177/000312240607100605

Beckfield, J. (2009). *Remapping Inequality in Europe: The Net Effect of Regional Integration on Total Income Inequality in the European Union*. International Journal of Comparative Sociology, 50(5–6), 486–509. doi:10.1177/0020715209339282

Blanchet, T., Chancel, L., & Gethin, A. (2019). *How unequal is Europe? evidence from distributional national accounts, 1980-2017.* World Inequality Database: World Inequality Lab. Retrieved from https://wid.world/document/bcg2019-full-paper/

Boschma, R. (2022). Regional diversification and inequality between and within regions. Retrieved from https://www.cesifo.org/DocDL/econpol-forum-2022-5-boschma-regional-diversification-inequality.pdf

Butkus, M., Mačiulytė-Šniukienė, A., Matuzevičiutė, K., & Cibulskienė, D. (2019). What is the return on investing in European regional development and cohesion funds? difference-in-differences estimator approach. *Ekonomický Časopis, 67*(6), 647-676. Retrieved from http://cejsh.icm.edu.pl/cejsh/element/bwmeta1.element.cejsh-7a8e208f-912a-4454-ae24-f5722d186c50

Butkus, M., Maciulyte-Sniukiene, A., & Matuzeviciute, K. (2020). Mediating effects of cohesion policy and institutional quality on convergence between EU regions: An examination based on a conditional beta-convergence model with a 3-way multiplicative term. *Sustainability, 12*, 3025. doi:10.3390/su12073025

Castells-Quintana, D., Ramos, R., & Royuela, V. (2015). Income inequality in european regions: Recent trends and determinants | SpringerLink. Retrieved from https://link.springer.com/article/10.1007/s10037-015-0098-4

Corporate Finance Institute. (2022). Solow growth model. Retrieved from https://corporatefinanceinstitute.com/resources/economics/solow-growth-model/

Delhey, J., & Dragolov, G. (2014). Why inequality makes Europeans less happy: The role of distrust, status anxiety, and perceived conflict. *European Sociological Review, 30*(2), 151-165. doi:10.1093/esr/jct033

Di Caro, P., & Fratesi, U. (2021). One policy, different effects: Estimating the region‐specific impacts of EU cohesion policy. *Journal of Regional Science, 62* doi:10.1111/jors.12566

Doran, J., & Jordan, D. (2013). Decomposing European NUTS2 regional inequality from 1980 to 2009. *Journal of Economic Studies (Bradford), 40*(1), 22-38. doi:10.1108/01443581311283484

European Commission. (2007). *Cohesion policy 2007–13: Commentaries and official texts*. Retrieved from https://www.lu.lv/materiali/biblioteka/es/pilnieteksti/ekonomika/Cohesion%20policy%202007-13.%20Commentaries%20and%20official%20texts.pdf

European Commission. (2015a). European structural and investment funds: 2014-2020 Retrieved from https://commission.europa.eu/funding-tenders/find-funding/funding-management-mode/2014-2020-european-structural-and-investment-funds_en

European Commission. (2015b). European structural and investment funds 2014-2020. Retrieved from https://op.europa.eu/en/publication-detail/-/publication/fb4e979d-b377-11e5-8d3c-01aa75ed71a1

European Commission. (n.a.a.). Reducing inequalities is essential to ensure sustainable development benefits all, especially those furthest behind. Retrieved from https://international-partnerships.ec.europa.eu/news-and-events/stories/reducing-inequalities-essential-ensure-sustainable-development-benefits-all-especially-those_en

European Commission. (n.a.b.). Inforegio - history of the policy. Retrieved from https://ec.europa.eu/regional_policy/policy/what/history_en

European Commission. (n.a.c.). Inforegio - is my region covered? (2007-2013). Retrieved from https://ec.europa.eu/regional_policy/policy/how/is-my-region-covered/2007-2013_en

European Commission. (n.a.d.). Inforegio - is my region covered? (2014-2020). Retrieved from https://ec.europa.eu/regional_policy/policy/how/is-my-region-covered/2014-2020_en

European Union. (n.a.a.). EU:S historia – 1945–1959. Retrieved from https://european-union.europa.eu/principles-countries-history/history-eu/1945-59_sv

European Union. (n.a.b.). Schuman declaration may 1950. Retrieved from https://european-union.europa.eu/principles-countries-history/history-eu/1945-59/schuman-declaration-may-1950_en

Eurostat. (2007). *Regions in the European union: Nomenclature of territorial units for statistics NUTS 2006 /EU-27*. Retrieved from https://ec.europa.eu/eurostat/documents/3859598/5902901/KS-RA-07-020-EN.PDF.pdf/5aceb8d2-f395-4843-92e9-a2dbd6443f16?t=1414781494000

Eurostat. (2015). *Regions in the European Union - 2013.* Retrieved from https://ec.europa.eu/eurostat/documents/3859598/6948381/KS-GQ-14-006-EN-N.pdf/b9ba3339-b121-4775-9991-d88e807628e3?t=1444229719000

Eurostat. (2021). Glossary: Income quintile share ratio. Retrieved from https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Glossary:Income_quintile_share_ratio

Eurostat. (2022). Statistics | eurostat - population by NUTS2 region. Retrieved from https://ec.europa.eu/eurostat/databrowser/view/DEMO_R_D2JAN__custom_4650666/default/table

Eurostat. (2023a). Statistics | eurostat - GDP per NUTS2 region. Retrieved from https://ec.europa.eu/eurostat/databrowser/view/nama_10r_2gdp/default/table?lang=en

Eurostat. (2023b). Statistics | eurostat - Employment by sex, age, job tenure and NUTS 2 regions. Retrieved from https://ec.europa.eu/eurostat/databrowser/view/LFST_R_EGAD__custom_5606381/default/table

Eurostat. (2023c). Statistics | eurostat - Pupils and students enrolled by education level, sex and NUTS2 regions. Retrieved from https://ec.europa.eu/eurostat/databrowser/view/EDUC_UOE_ENRA11__custom_5625835/default/table

Eurostat. (2023d). Statistics | eurostat - Employment in technology and knowledge-intensive sectors by NUTS 2 regions and sex. Retrieved from https://ec.europa.eu/eurostat/databrowser/view/HTEC_EMP_REG2__custom_5625928/default/table

Eurostat. (2023e). Statistics | eurostat - Employment rates by sex, age, educational attainment level, citizenship and NUTS 2 regions. Retrieved from https://ec.europa.eu/eurostat/databrowser/view/LFST_R_LFE2EMPRTN__custom_5625825/default/table

Eurostat. (2023f). Statistics | eurostat - Population density by NUTS 2 region. Retrieved from https://ec.europa.eu/eurostat/databrowser/view/TGS00024__custom_5625897/default/table

Eurostat. (2023g). Statistics | eurostat - Population on 1 January by age, sex and NUTS 2 region. Retrieved from https://ec.europa.eu/eurostat/databrowser/view/DEMO_R_D2JAN__custom_5625908/default/table

Eurostat. (2023h). Statistics | eurostat - R&D personnel and researchers by sector of performance, sex and NUTS 2 regions. Retrieved from https://ec.europa.eu/eurostat/databrowser/view/RD_P_PERSREG__custom_5625869/default/table

Eurostat. (n.a.a.). Background - NUTS - nomenclature of territorial units for statistics - eurostat. Retrieved from https://ec.europa.eu/eurostat/web/nuts/background

Eurostat. (n.a.b.). Statistics | eurostat. Retrieved from https://ec.europa.eu/eurostat/databrowser/view/ILC_DI11_R__custom_4614936/default/table?lang=en

Eurostat. (n.a.c.). Principles and characteristics. Retrieved from https://ec.europa.eu/eurostat/en/web/nuts/principles-and-characteristics

Ezcurra, R. (2019). Regional disparities and within-country inequality in the European Union. *Revista De Economía Mundial, 51*(51) doi:10.33776/rem.v0i51.3907

Farole, T., Rodríguez-pose, A., & Storper, M. (2011). Cohesion policy in the European union: Growth, geography, institutions. *Journal of Common Market Studies, 49*, 1089-1111. doi:10.1111/j.1468-5965.2010.02161.x

Gabel, M. (2022). European union - creation of the european economic community | britannica. Retrieved from https://www.britannica.com/topic/European-Union

Gründler, K., & Scheuermeyer, P. (2018). Growth effects of inequality and redistribution: What are the transmission channels? *Journal of Macroeconomics, 55*, 293-313. Retrieved from https://www.sciencedirect.com/science/article/pii/S0164070416302336?via%3Dihub

Heiss, A. (2020). Fuzzy regression discontinuity. Retrieved from https://evalf20.classes.andrewheiss.com/example/rdd-fuzzy/

Holla, A. (n.a.). Regression discontinuity. Retrieved from https://thedocs.worldbank.org/en/doc/554591525386187332-0090022018/original/16.TechnicalTrackRegressionDiscontinuity.pdf

IMF. (n.a.). Introduction to inequality. Retrieved from https://www.imf.org/en/Topics/Inequality/introduction-to-inequality

Jianu, I., Dinu, M., Huru, D., & Bodislav, A. (2021). Examining the relationship between income inequality and growth from the perspective of EU member states' stage of development. *Sustainability (Basel, Switzerland), 13*(9), 5204. doi:10.3390/su13095204

Kersan-Škabić, I., & Tijanić, L. (2017). Regional absorption capacity of EU funds. *Economic Research-Ekonomska Istraživanja, 30*(1), 1191-1208. doi:10.1080/1331677X.2017.1340174

Khun, T., Elsas, E., Hakhverdian, A., & Brug, W. (2014). An ever wider gap in an ever closer union: Rising inequalities and euroscepticism in 12 west european democracies, 1975–2009. *Socio-Economic Review, 14*(1), 27-45. doi:10.1093/ser/mwu034

Koudoumakis, P., Botzoris, G., & Protopapas, A. (2021). The contribution of cohesion policy to the development and convergence of the regions of the European Union. *Regional Science Inquiry, XIII*(2), 277-290. Retrieved from https://ideas.repec.org//a/hrs/journl/vxiiiy2021i2p277-290.html

Kuznets, S. (1955). Economic growth and income inequality. *The American Economic Review, XLV* Retrieved from https://assets.aeaweb.org/asset-server/files/9438.pdf

Kyriacou, A., Maynou L., Saez M., & Bacaria, J. (2014). The impact of structural and cohesion funds on eurozone convergence (1990-2010). *Regional Studies, 50* doi:10.1080/00343404.2014.965137

Kyriacou, A., & Roca-Sagalés, O. (2012). The impact of EU structural funds on regional disparities within member states. *Environment and Planning C: Government and Policy, 30*, 267-281. doi:10.1068/c11140r

Lansley, S. (2012). Inequality and instability: Why more equal societies have more stable economies. Retrieved from https://cpag.org.uk/sites/default/files/CPAG-Poverty142-inequality-instability.pdf

Malikov, A., & Alimov, B. (2022). Income inequality and political instability. Retrieved from https://iariw.org/wp-content/uploads/2022/08/Malikov-Alimov-IARIW-2022.pdf

McCrary, J. (2006). *Manipulation of the running variable in the regression discontinuity design: A density test**. Retrieved from https://eml.berkeley.edu/~jmccrary/mccrary2006_DCdensity.pdf

Medve-Bálint, G. (2018). The cohesion policy on the EU's eastern and southern periphery: Misallocated funds? *Studies in Comparative International Development,* Retrieved from https://link.springer.com/article/10.1007/s12116-018-9265-2

Mohl, P., & Hagen, T. (2010). Do EU structural funds promote regional growth? new evidence from various panel data approaches. *Regional Science and Urban Economics, 40*(5), 353-365. doi:10.1016/j.regsciurbeco.2010.03.005

Naughton, B. (1993). Deng xiaoping: The economist. *The China Quarterly*, 501. Retrieved from https://www.jstor.org/stable/654099?seq=3

Neves, P. C., Afonso, Ó, & Silva, S. T. (2016). A meta-analytic reassessment of the effects of inequality on growth. *World Development, 78*, 386-400. Retrieved from https://www.sciencedirect.com/science/article/pii/S0305750X15002600?via%3Dihub

Nordregio. (n.a.). EU cohesion policy 2007-2013. Retrieved from https://archive.nordregio.se/en/Metameny/About-Nordregio/Journal-of-Nordregio/2006/Journal-of-Nordregio-no-2-2006/EU-cohesion-policy-2007-2013/index.html

Papatheodorou, C., & Pavlopoulos, D. (2013). Income inequality in the EU: How do member states contribute? Retrieved from https://www.emerald.com/insight/content/doi/10.1108/IJSE-02-2013-0030/full/html

Pinho, C., Varum, C., & Antunes, M. (2015). Under what conditions do structural funds play a significant role in European regional economic growth? some evidence from recent panel data. *Journal of Economic Issues, 49*, 749-771. doi:10.1080/00213624.2015.1072382

Romer, P. M. (1990). Endogenous technological change. *The Journal of Political Economy, 98*(5) Retrieved from http://web.stanford.edu/~klenow/Romer_1990.pdf

Satz, D., & White, S. (2021). What is wrong with inequality? Retrieved from https://ifs.org.uk/inequality/what-is-wrong-with-inequality/

Schraff, D., & Lipps, J. (2020). Regional inequality and institutional trust in europe. *European Journal of Political Research, 60* doi:10.1111/1475-6765.12430

Shin, I. (2012). Income inequality and economic growth. *Economic Modelling, 29*(5), 2049-2057. Retrieved from https://www.sciencedirect.com/science/article/pii/S0264999312000466

Soava, G., Mehedintu, A., & Sterpu, M. (2019). Relations between income inequality, economic growth and poverty threshold: New evidence from EU countries panels. *Technological and Economic Development of Economy, 26*, 1-21. doi:10.3846/tede.2019.11335

Staehr, K., & Urke, K. (2022). The European structural and investment funds and public investment in the EU countries. Retrieved from https://link.springer.com/article/10.1007/s10663-022-09549-6

Stanley, A. (2022). Global inequalities. Retrieved from https://www.imf.org/en/Publications/fandd/issues/2022/03/Global-inequalities-Stanley

Stewart, B. (2016). Causality with unmeasured confounding. Retrieved from https://scholar.princeton.edu/sites/default/files/bstewart/files/lecture11handout.pdf

Stock, H. J., & Watson, W. M. (2020). *Introduction to econometrics* (Fourth ed.) Pearson.

United Nations. (2017). *Equality: Why it matters.* Retrieved from https://www.un.org/sustainabledevelopment/wp-content/uploads/2017/02/ENGLISH_Why_it_Matters_Goal_10_Equality.pdf

Venables, A. J. (2016). New economic geography. Retrieved from https://link.springer.com/referenceworkentry/10.1057/978-1-349-95121-5_2329-1#Sec1

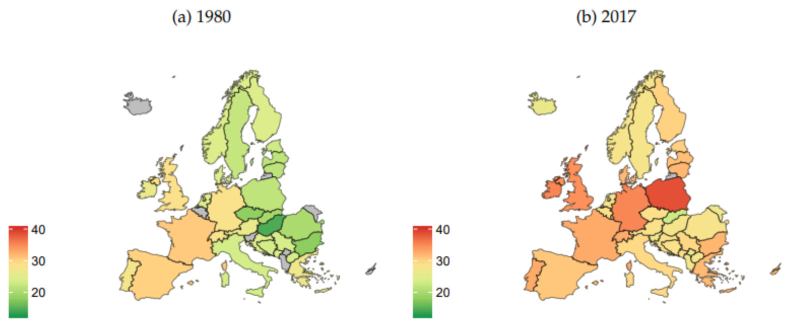
Vukašina, M., Kersan-Škabiæ, I., & Orliæ, E. (2022). Impact of European structural and investment funds absorption on the regional development in the EU–12 (new member states). *Equilibrium. Quarterly Journal of Economics and Economic Policy, 17*(4), 857-880. Retrieved from https://ideas.repec.org//a/pes/ierequ/v17y2022i4p857-880.html

Wing, C., & Bello-Gomez, R. A. (2018). Regression discontinuity and beyond: Options for studying external validity in an internally valid design. *American Journal of Evaluation, 39*(1), 91-108. doi:10.1177/1098214017736155

Appendix

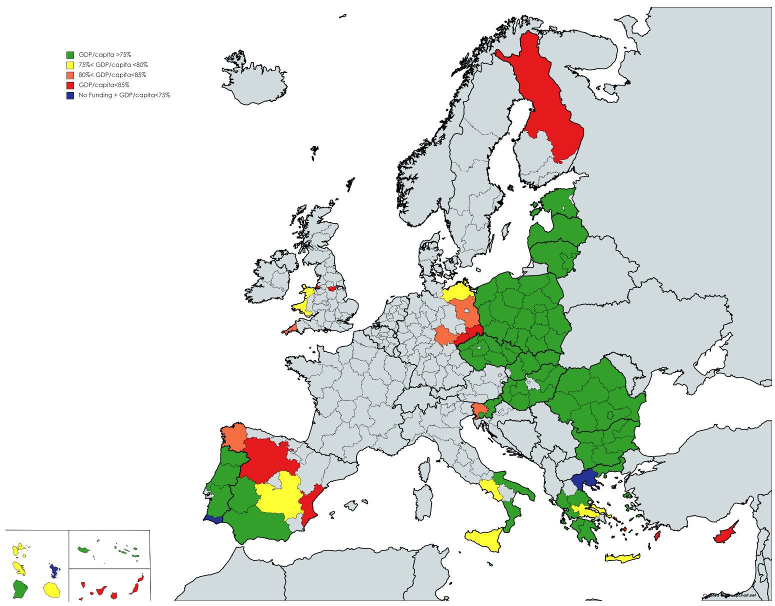
A. Figures

Figure 1: Top 10% income earner's share of national income



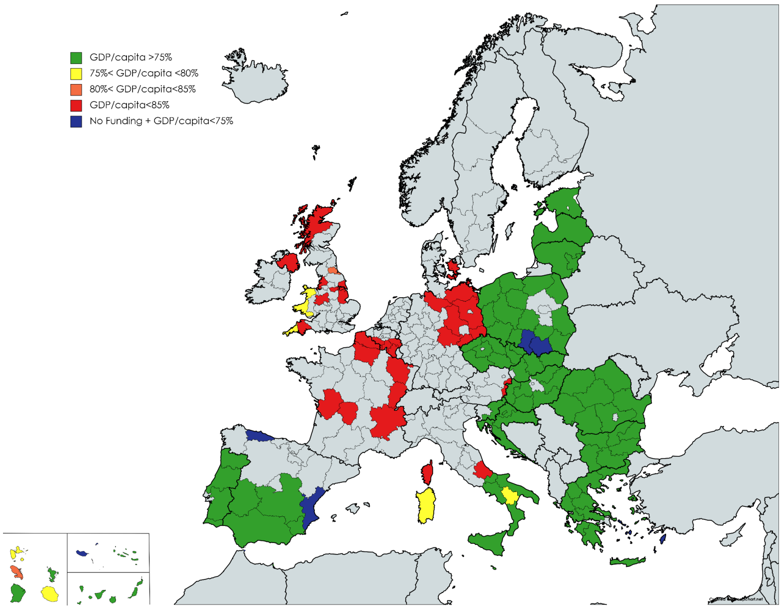
Source: (Blanchet et al., 2019)

Figure 2: Figure representing regions receiving funding between 2007-2013



Source: (European Commission, 2007)

Figure 3: Figure representing regions receiving funding between 2014-2020



Source: (European Commission, 2015b)

B. Tables

Table 3: Descriptive statistics over employment, employment per thousands of persons

Country	Average	Min	Max
Austria	490	131	914
Belgium	412	105	818
Bulgaria	508	254	1050
Croatia	408	379	425
Cyprus	374	350	405
Czech Republic	628	487	817
Denmark	533	256	927
Estonia	603	548	632
Finland	486	14	831
France	922	43	5406

Germany	1010	246	2506
Greece	296	61	1761
Hungary	458	324	632
Ireland	897	314	1815
Italy	1060	53	4372
Latvia	881	829	1016
Luxembourg	244	202	290
Malta	192	155	256
Netherlands	689	177	1846
Poland	861	342	1885
Portugal	642	96	1681
Romania	1049	707	1545
Slovakia	601	307	890
Slovenia	467	425	514
Spain	979	27	3559
Sweden	579	162	1232
UK	690	216	1300

Source: (Eurostat, 2023b)

Table 4: Descriptive statistics over income inequality

Country	Average	Min	Max
Bulgaria	6,6	4,6	9,4
Croatia	4,7	4,4	5,1
Cyprus	4,6	4,3	5,4
Denmark	4,0	2,9	6,0
Estonia	5,4	5,0	6,5
Finland	3,5	3,1	4,1
Hungary	3,8	2,8	5,1
Ireland	4,2	3,5	5,1
Italy	5,2	3,4	10,4
Latvia	6,6	6,2	7,4
Luxembourg	4,5	4,0	5,3
Malta	4,2	3,9	4,7
Romania	6,6	4,0	11,6
Slovakia	3,5	2,6	4,9
Slovenia	3,4	3,2	3,7
Sweden	3,9	3,0	8,4

*Source: (*Eurostat, 2021)

Table 5: Regions used for income inequality analysis
Bulgaria: Severozapaden, Severen tsentralen, Severoiztochen, Yugoiztochen, Yugozapaden, Yuzhen tsentralen

- Due to missing values for all regions in Bulgaria in 2007, we have extrapolated the income quintile share ratio from the year 2008.

Denmark: Hovedstaden, Sjælland, Syddanmark, Midtjylland, Nordjylland
Estonia
Ireland:* Border Midland and Western, Southern and Eastern (up until 2013), Northern and Western, Southern, Eastern and Midland (after 2013).

- In Ireland, the NUTS 2 regions have changed between 2007 to 2020. Previously, there were only 2 NUTS 2 regions but in 2016, there was a revision when a third NUTS 2 region was created. Data for the previous NUTS 2 regions only exist up until 2011. We will assume that the original region Southern and Eastern is equivalent to the regions Southern and to Eastern and Midland while the original region Border, Midland and Western is equivalent to Northern and Western. We have also extrapolated data for 2014 and 2015 for the three new regions. The 2014 data comes from the 2013 data and the 2015 from the 2016 data.

Croatia: Jadranska Hrvatska (only part of second programming period since they joined the EU in 2013)

Italy: Piemonte, Valle d'Aosta/Vallée d'Aoste, Liguria, Lombardia, Provincia Autonoma di, Bolzano/Bozen, Provincia Autonoma di Trento, Veneto, Friuli-Venezia Giulia, Emilia-Romagna, Toscana, Umbria, Marche, Lazio, Abruzzo, Molise, Campania, Puglia, Basilicata, Calabria, Sicilia, Sardegna

Cyprus

Latvia

Luxembourg

Hungary: Közép-Magyarország, Közép-Dunántúl, Nyugat-Dunántúl, Dél-Dunántúl, Észak-Magyarország, Észak-Alföld, Dél-Alföld

- The capital is located in the NUTS 2 region Közép-Magyarország. However, due to large differences in wealth between the habitants in the inner city and the habitants in the suburb, predominantly on the Pest side the NUTS 2 region has now been split into 2 areas. However, there is only data for these two areas from 2018 onwards. Thus, we will neglect this.

Malta

Romania: Nord-Vest, Centru, Nord-Est, Sud-Est, Sud - Muntenia, București - Ilfov, Sud-Vest Oltenia, Vest

Slovenia: Vzhodna, Zahodna

- Extrapolated value for 2007 from 2008.

Slovakia: Bratislavský kraj, Západné, Stredné, Východné

Finland: Itä-Suomi, Etelä-Suomi, Pohjois-Suomi (up until 2006), Länsi-Suomi, Åland, Helsinki-Uusimaa, Etelä-Suomi, Pohjois- ja Itä-Suomi

- In Finland the original NUTS 2 regions have been changed similarly to Ireland. Data for the new NUTS 2 regions is missing and thus we assume that the two original regions Itä-Suomi and Pohjois-Suomi are equivalent to Pohjois-ja Itä-Suomi. We also assume that the original region Etelä-Suomi is equivalent to Helsinki-Uusimaa and Etelä-Suomi.

Sweden: Stockholm, Östra Mellansverige, Småland med öarna, Sydsverige, Västsverige, Norra Mellansverige, Mellersta Norrland, Övre Norrland

- Extrapolated values for 2007 from 2008.

*Source: (*European Commission, 2007; European Commission, 2015b)

Table 6: Regions used for employment analysis which are matched to the income inequality dataset
Same regions as in table 4 with the exception of Slovenia as there was no complete data over the whole time period, and was hence excluded.
*Source: (*European Commission, 2007; European Commission, 2015b)

Table 7: Regions used for the full employment analysis
Austria: Burgenland, Kärnten, Niederösterreich, Oberösterreich, Salzburg, Steiermark, Tirol, Vorarlberg, Wien
Belgium: Prov. Antwerpen, Prov. Brabant wallon, Prov. Hainaut, Prov. Liège, Prov. Limburg (BE), Prov. Luxembourg (BE), Prov. Namur, Prov. Oost-Vlaanderen, Prov. Vlaams-Brabant, Prov. West-Vlaanderen, Région de Bruxelles-Capitale/Brussels Hoofdstedelijk Gewest
Bulgaria: Severen tsentralen, Severoiztochen, Severozapaden, Yugoiztochen, Yugozapaden, Yuzhen tsentralen
Croatia: Sostines regionas
Cyprus
Czech Republic: Jihovýchod, Jihozápad, Praha, Severovýchod, Severozápad, Střední Čechy, Střední Morava
Denmark: Hovedstaden, Midtjylland, Nordjylland, Sjælland, Syddanmark
Estonia
Finland: Etelä-Suomi, Helsinki-Uusimaa, Länsi-Suomi, Pohjois- ja Itä-Suomi, Åland
France: Alsace, Aquitaine, Auvergne, Basse-Normandie, Bourgogne, Bretagne, Centre - Val de Loire, Champagne-Ardenne, Corse, Franche-Comté, Guadeloupe, Guyane, Haute-Normandie, Île de France, La Réunion, Languedoc-Roussillon, Limousin, Lorraine, Martinique, Midi-Pyrénées, Nord-Pas-de-Calais, Pays-de-la-Loire, Picardie, Poitou-Charentes, Provence-Alpes-Côte d'Azur, Rhône-Alpes, Mayotte
Germany: Arnsberg, Brandenburg, Braunschweig, Bremen, Chemnitz, Darmstadt, Detmold, Dresden, Düsseldorf, Freiburg, Gießen, Hamburg, Hannover, Karlsruhe, Kassel, Koblenz, Köln, Leipzig, Lüneburg, Mecklenburg-Vorpommern, Mittelfranken, Münster, Niederbayern, Oberbayern, Oberfranken, Oberpfalz, Rheinhessen-Pfalz, Saarland, Sachsen-Anhalt, Schleswig-Holstein, Schwaben, Stuttgart, Thüringen, Trier, Tübingen, Unterfranken, Weser-Ems, Berlin
Greece: Anatoliki Makedonia Thraki, Attiki, Dytiki Ellada, Dytiki Makedonia, Ionia Nisia, Ipeiros, Kentriki Makedonia, Kriti, Notio Aigaio, Peloponnisos, Sterea Ellada, Thessalia, Voreio Aigaio
Hungary: Dél-Alföld, Dél-Dunántúl, Észak-Alföld, Észak-Magyarország, Közép-Dunántúl, Nyugat-Dunántúl
Ireland: Eastern and Midland, Northern and Western, Southern
Italy: Abruzzo, Basilicata, Calabria, Campania, Emilia-Romagna, Friuli-Venezia, Giulia, Lazio, Liguria, Lombardia, Marche, Molise, Piemonte, Provincia Autonoma di Bolzano/Bozen, Provincia Autonoma di Trento, Puglia, Sardegna, Sicilia, Toscana, Umbria, Valle d'Aosta/Vallée d'Aoste, Veneto
Latvia

Lithuania: Vidurio ir vakaru
Luxembourg
Malta
Netherlands: Drenthe, Flevoland, Friesland (NL), Gelderland, Groningen, Limburg (NL), Noord-Brabant, Noord-Holland, Overijssel, Utrecht, Zeeland, Zuid-Holland
Poland: Dolnoslaskie, Kujawsko-Pomorskie, Łódzkie, Lubelskie, Lubuskie, Małopolskie, Moravskoslezsko, Opolskie, Podkarpackie, Podlaskie, Pomorskie, Śląskie, Swietokrzyskie, Warminsko-Mazurskie, Wielkopolskie, Zachodniopomorskie
Portugal: Alentejo, Algarve, Área Metropolitana de Lisboa, Centro (PT), Norte, Região Autónoma da Madeira (PT), Região Autónoma dos Açores (PT)
Romania: Bucuresti - Ilfov, Centru, Nord-Est, Nord-Vest, Sud-Est, Sud-Vest Oltenia, Sud - Muntenia, Vest
Slovakia: Bratislavský kraj, Stredné Slovensko, Východné Slovensko, Západné Slovensko
Slovenia: Vzhodna Slovenija, Zahodna Slovenija
Spain: Andalucía, Aragón, Canarias, Cantabria, Castilla-la Mancha, Castilla y León, Cataluña, Ciudad de Ceuta, Ciudad de Melilla, Comunidad de Madrid, Comunidad Foral de Navarra, Comunitat Valenciana, Extremadura, Galicia, Illes Balears, La Rioja, País Vasco, Principado de Asturias, Región de Murcia
Sweden: Mellersta Norrland, Norra Mellansverige, Småland med öarna, Stockholm, Sydsverige, Västsverige, Östra Mellansverige, Övre Norrland
United Kingdom: Bedfordshire and Hertfordshire, Berkshire Buckinghamshire and Oxfordshire, Cheshire, Cornwall and Isles of Scilly, Cumbria, Derbyshire and Nottinghamshire, Devon, Dorset and Somerset, East Anglia, East Wales, East Yorkshire and Northern Lincolnshire, Essex, Gloucestershire Wiltshire and Bristol/Bath area, Greater Manchester, Hampshire and Isle of Wight, Herefordshire Worcestershire and Warwickshire, Highlands and Islands, Kent, Lancashire, Leicestershire Rutland and Northamptonshire, Lincolnshire, Merseyside, North Eastern Scotland, North Yorkshire, Northumberland and Tyne and Wear, Shropshire and Staffordshire, South Yorkshire, Surrey East and West Sussex, Tees Valley and Durham, West Midlands, West Wales and The Valleys, West Yorkshire.
*Source: (*European Commission, 2007; European Commission, 2015b)

Table 14: Income inequality analysis neglecting phasing out regions

Bandwidths	None	+/- 70%	+/- 50%	+/- 20%
Intercept	0.10	0.30	0.63	-0.67
	(0.06)	(0.26)	(0.81)	(1.62)
GDP	-0.07	-0.50	-1.43	3.54
	(0.06)	(0.47)	(1.98)	(6.52)
Treatment	-0.17	-0.57	-1.21	1.48
	(0.09)	(0.48)	(1.61)	(3.12)
Observations	146	107	70	26
p-value *of treatment*	0.08	0.24	0.45	0.64

*Notes: *** p < 0.001; ** p < 0.01; * p < 0.05*
Sources: (Eurostat, 2021; Eurostat, 2022; Eurostat, 2023a; European Commission 2015b; European Commission 2007)

Table 15: Employment analysis using matched dataset and neglecting phasing-out regions

Bandwidths	None	+/- 0.7	+/- 0.5	+/- 0.2
Intercept	-5.34	-12.45	-5.49	41.06
	(2.91)	(15.05)	(29.7)	(86.92)
GDP	7.42*	21.8	3.72	-158.92
	(3.23)	(30.62)	(79.31)	(363.17)
Treatment	5.38	19.37	6.44	-94.45
	(5.82)	(31.02)	(63.83)	(174.32)
Observations	142	103	66	22
p-value *of treatment*	0.36	0.53	0.92	0.72

Notes: *** p < 0.001; ** p < 0.01; * p < 0.05
Sources: (Eurostat, 2022; Eurostat, 2023a; Eurostat, 2023b; European Commission, 2015b; European Commission, 2007)

Table 16: Employment analysis using full dataset and neglecting phasing-out regions

Bandwidths	None	+/- 0.7	+/- 0.5	+/- 0.2
Intercept	-9.16**	-32.09	-37.99	-90.61
	(3.42)	(21.59)	(43.33)	(285.1)
GDP	17.79***	57.78	69.43	206.93
	(4.09)	(37.25)	(82.04)	(720.57)
Treatment	15.11**	24.44	63.27	129.1
	(5.37)	(36.42)	(71.58)	(416.21)
Observations	500	414	325	109
p-value *of treatment*	0.003	0.08	0.38	0.31

Notes: *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$
Sources: (Eurostat, 2022; Eurostat, 2023a; Eurostat, 2023b; European Commission, 2015b; European Commission, 2007)

Table 17: Regions for robustness checks
When doing the robustness checks, we aimed at using the same regions and same time periods as in table 4 and table 7, however due to data limitations this was not fully possible for all our robustness checks. Thus, these are the modifications for our robustness checks made for income inequality:

- **Pupils and students enrolled in short-cycle tertiary education**: Severozapaden, Severen tsentralen, Severoiztochen, Yugoiztochen, Yugozapaden, Yuzhen tsentralen, Estonia, Nord-Vest, Centru, Nord-Est, Sud-Est, Sud - Muntenia, București - Ilfov, Sud-Vest Oltenia, Vest were removed. Additionally, data was only existing for the years 2013-2020. Thus, programming period 1 will solely be made up of data from 2013. There were no data for the regions in Ireland in 2013, which implies that those regions will be excluded.
- **Employment in technology and knowledge-intensive sectors**: Data missing for year 2007 and the region Valle d'Aosta/Vallée d'Aoste.
- **Employment rates:** No modifications
- **Population density:** No data for the years 2007 and 2020.
- **Population level:** No modifications
- **R&D personnel:** No modifications

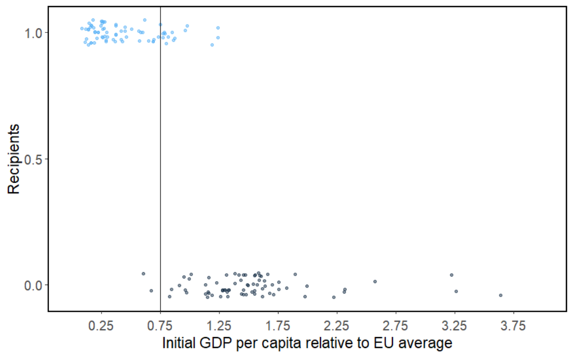
When doing the robustness checks for employment, we also had to make some modifications. Due to the large amount of regions in this dataset, we will only list the countries, and not all the regions, which we lack data for when doing the robustness checks:

- **Pupils and students enrolled in short-cycle tertiary education**: Data was only available for 2013-2020. Countries missing: Bulgaria, Finland, Germany, Greece, Lithuania, Romania and Great Britain.
- **Employment in technology and knowledge-intensive sectors**: Data missing for year 2007. No missing countries.
- **Employment rates:** We will not use employment rate as robustness check as this data is very similar to employment level.
- **Population density:** Missing data for 2007 and 2020. No missing countries.
- **Population level:** No modifications.
- **R&D personnel:** Missing data for France and the Netherlands in the second programming period.

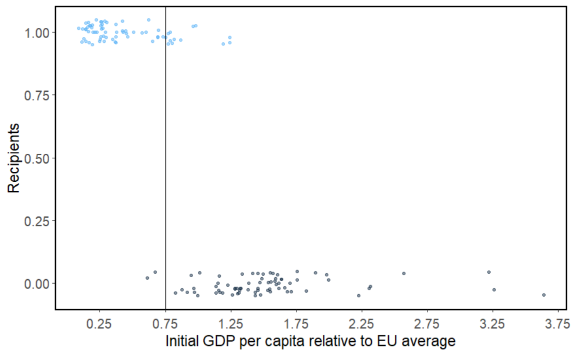
Sources: (Eurostat, 2022; Eurostat, 2023a; Eurostat, 2023b; Eurostat, 2023c; Eurostat, 2023d; Eurostat, 2023e; Eurostat, 2023f)

C. Graphs

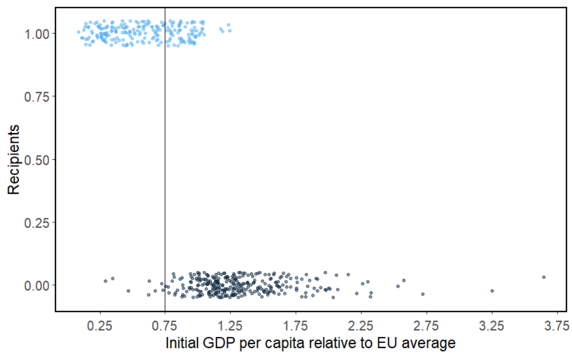
Graph 1: Compliance to the threshold rule



Compliance for data set measuring income inequality



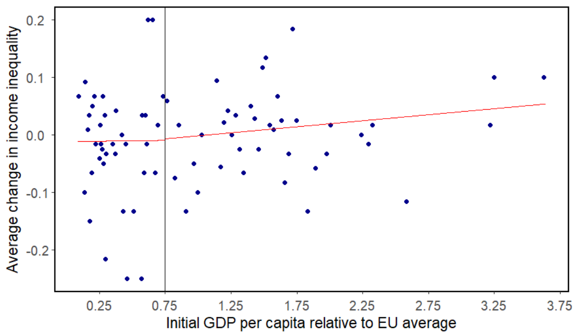
Compliance for data set measuring employment - matched



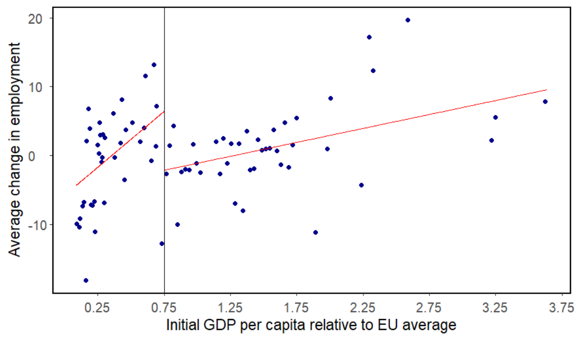
Compliance for dataset measuring employment - matched

Sources: (Eurostat, 2022; Eurostat, 2023a; European Commission, 2007; European Commission, 2015b)

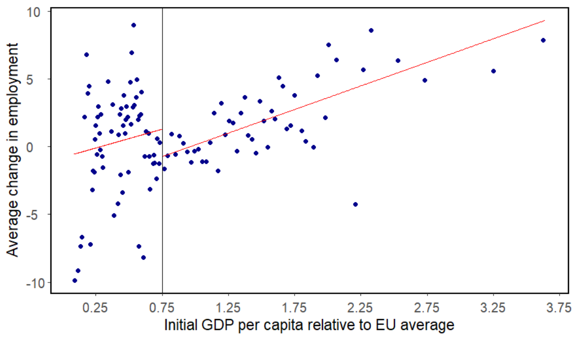
Graph 3: Removing outliers - main results



Average change in income inequality



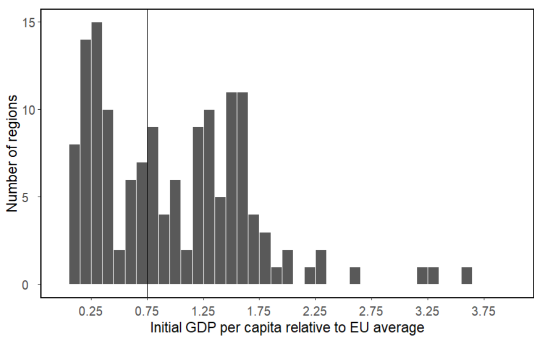
Average change in employment - matched



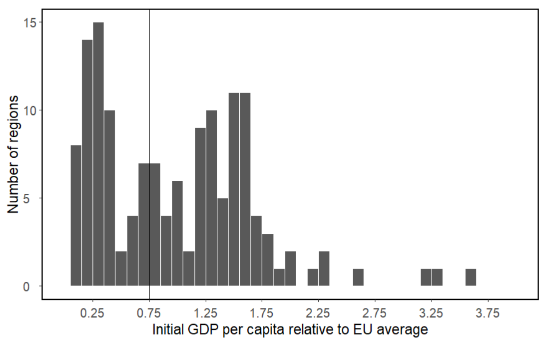
Average change in employment - full data

Sources: (Eurostat, 2021; Eurostat, 2022; Eurostat, 2023a; Eurostat, 2023b; European Commission, 2015b; European Commission, 2007)

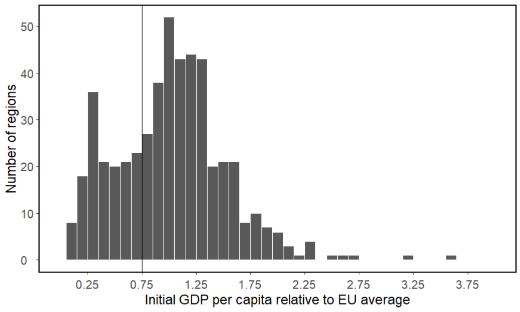
Graph 6: Distribution of regions



Distribution of regions across different levels of GDP/capita for the data set measuring income inequality.



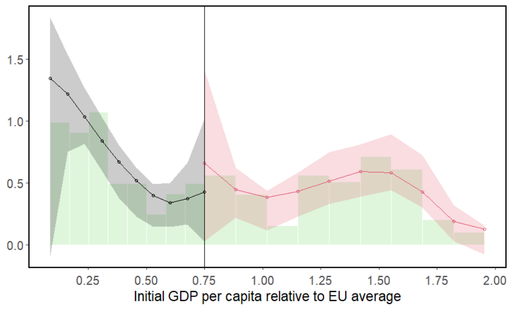
Distribution of regions across different levels of GDP/capita for the data set measuring employment - matched.



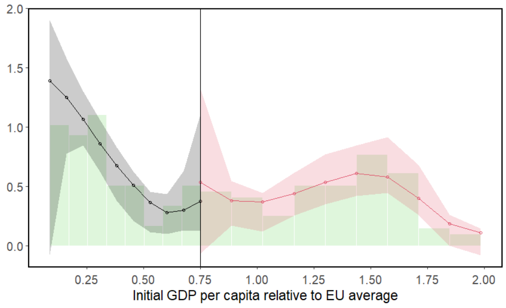
Distribution of regions across different levels of GDP/capita for the dataset measuring employment - full.

Sources: (Eurostat, 2022; Eurostat 2023a)

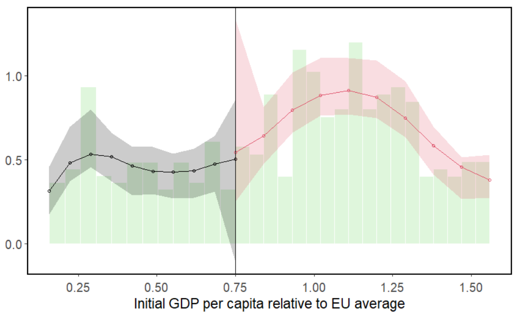
Graph 7: Density testing



Density plot of the income inequality dataset



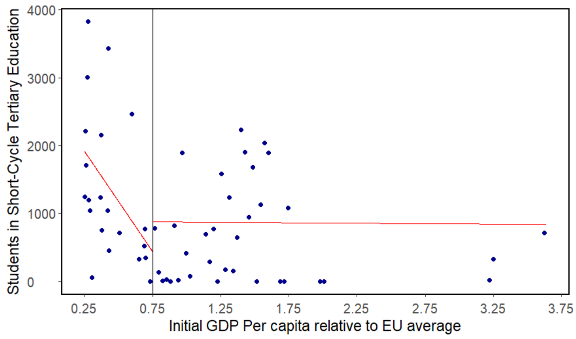
Density plot of the employment - matched dataset



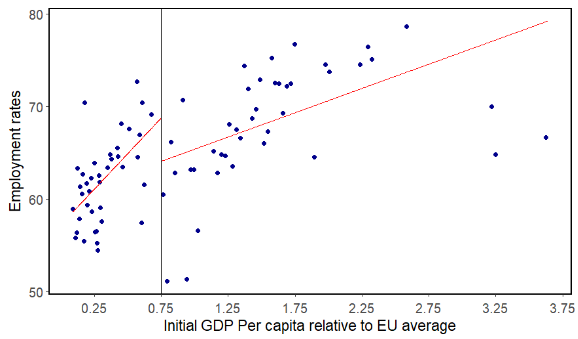
Density plot of the employment - full dataset

Sources: (Eurostat, 2021; Eurostat, 2022; Eurostat, 2023a; Eurostat, 2023b; European Commission, 2015b; European Commission, 2007)

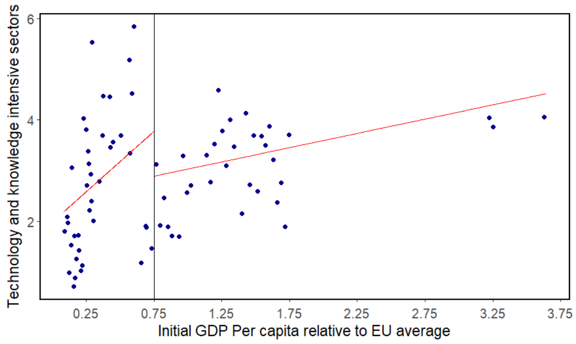
Graph 9: Using regions matched to the income inequality analysis without outliers



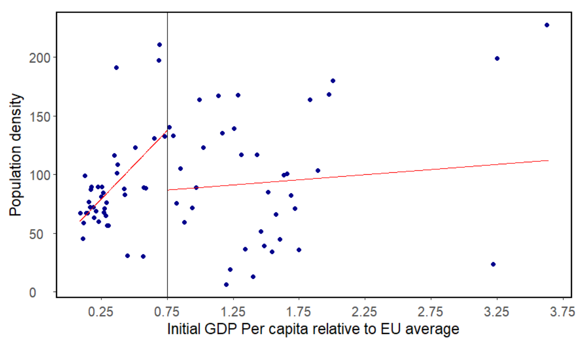
Students in Short- Cycle Tertiary Education



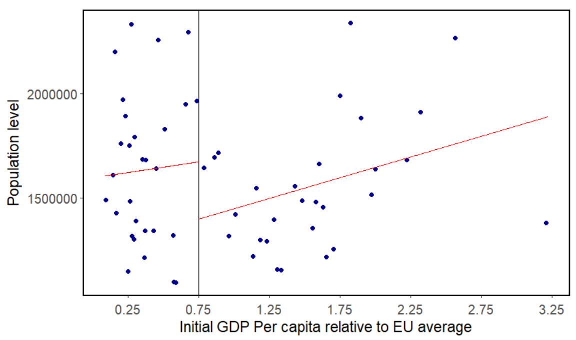
Employment rate



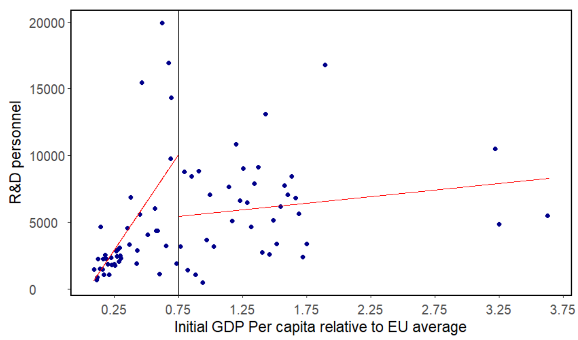
Employment in technology and knowledge intensive sectors



Population density



Population level



R&D personnel

Sources: (Eurostat, 2022; Eurostat, 2023a; Eurostat, 2023b; Eurostat, 2023c; Eurostat, 2023d; Eurostat, 2023e; Eurostat, 2023f)