# EU cohesion policy: in the name of economic growth?

Veiga da Silva, Gonçalo

**LEIDEN UNIVERSITY**

MSc Public Administration

Specialisation Economics and Governance

# EU cohesion policy: in the name of economic growth?

Assessing the impact of ´Investment for growth and jobs` funds in less developed regions in 2014-2020

## Master´s Thesis in Public Administration

*Prof. Eduard Suari Andreu*

Gonçalo Veiga da Silva (s2999277)

09/03/2022

# Abstract

The European Union has an ambitious cohesion policy, which aims to promote the economic growth and social development of its less developed territories. In order to achieve this goal, the EU relies on the European Structural and Investment Funds, which provide the necessary funding for cohesion actions. In the programming period 2014-2020, the EU defined an ´Investment for growth and jobs` goal, under which the biggest slice of those funds was channeled to NUTS 2 regions that complied with the rule of having a GDP per capita below 75% of the EU average (less developed regions). The application of the ´75%` criterion suits the application of a regression discontinuity design to assess the effect of the funds. Using this methodology, this Master´s thesis finds a positive significant effect of the funds on the economic growth of less developed regions during the programming period 2014 and 2020. The findings also suggest that the funds had an equal impact across different geographical groups of regions.

# Acknowledgments

*By the time I am concluding this study,*

*I am genuinely grateful, and I want to acknowledge that:*

To my parents and grandparents for their love and unconditional support in every moment of my life, including during my studies abroad.
To my brother, for his companionship for a lifetime, and for making me laugh.

To Rita, for the deepest partnership one can have,
for the times she simplified the world for me, and for her strength.

To my supervisor, for guiding me in the process of writing this thesis, for the constant availability, and for the understanding of my highs and lows when studying this topic.

And to all those I had the privilege to cross during my Master´s studies
and which were part of this journey and the process of developing this thesis.

# TABLE OF CONTENTS

# 1.
# Introduction

Economic and social policies in the European Union (EU) are multiple and aim at accomplishing different goals. Among them, the promotion of cohesion between Member-States and regions has long attained particular attention and even deserved consecration at the highest legal level of the Union. As today´s Article 174 of the Treaty on the Functioning of the European Union (2009) poses: "the Union shall develop and pursue its actions leading to the strengthening of its economic, social and territorial cohesion" and "shall aim at reducing disparities between the levels of development of the various regions and the backwardness of the least favored regions.".

To deliver social and economic growth to its less developed regions, and consequently to foster their convergence with the wealthiest ones, the European Union established its own cohesion policy which, throughout the years, has consecutively acquired more and more relevance both in political and budgetary terms (Rodriguez-Pose and Fratesi, 2004). Hence, after incipient forms of a common regional policy, the EU would build a way more ambitious scheme to tackle regional economic disparities, which eventually became today´s called ´EU Cohesion Policy`[1]. In fact, since 2007 that the cohesion policy started to represent the main expenditure item of the EU´s budget, replacing the well-known agricultural actions (Benedetto, 2019). Nowadays, it accounts for 25% of the EU´s multiannual financial framework for 2021-2027 (Benedetto, 2019), a value translated into around 392 billion euros (European Commission, n.d.).

The application of the available budget for the cohesion policy mainly relies on the provision of different financial schemes which are part of the today named

[1] A more complete picture of the development of the EU´s Cohesion Policy will be drawn in section 2.1.

´European Structural and Investment Funds` (hereinafter also ´ESIF`) (European Commission, n.d.). The implementation of the ESIF – including those devoted to the cohesion policy - is guided by policy goals that are previously defined. Under these policy goals, the funds are delivered to recipient Member-States and regions. Both the policy goals are designed and the funds are implemented in specific time windows, the so-called ´programming periods` of the ESIF. In order to be granted with funds, Member-States and regions usually have to fulfill certain previously established criteria.

Over the years, the main policy goal of the ESIF has also been the one of the cohesion policy. Despite its different formulations during the distinct programming periods, this main policy goal has been focused on promoting convergence in the EU and delivering greater economic growth and a raise of income per capita levels of the less developed regions (European Commission, 2007; European Commission, 2015). Since the major reform of the ESIF in 1988 that most of the ESIF have been allocated exactly to this policy goal and, under that, to the less developed regions of the EU. Illustratively, during the programming period for the round of 2007 to 2013, the less developed regions were allocated 57,5% of the total amount of funds devoted to the cohesion policy (European Commission, 2007). In turn, in the last programming period (from 2014 to 2020), that value was around 52% (European Commission, 2015).

The attempts to evaluate whether the EU´s cohesion policy is delivering the desired outcomes have been multiple. In particular, the research on the effects of ESIF in less developed regions has so far reached dissimilar conclusions as to whether the EU transfers are really promoting GDP per capita growth and fostering the convergence of these regions with the more developed territories (Boldrin and Canova, 2001; Esposti and Bussoletti, 2008; Becker et al., 2010). A potential explanation for this lack of consensus lies in the multitude of empirical strategies, datasets, and goals of the research conducted so far (Esposti and Bussoletti, 2008; Hagen and Mohl, 2009; Crescenzi and Giua, 2020).

Nevertheless, the evaluation of the effects of the EU cohesion policy - and, in particular, of the applications of ESIF in the less developed regions - remains highly relevant. This assessment is crucial for a more sophisticated and effective future policy design and implementation. If we agree that promoting economic growth and increasing income levels in all its countries and regions is a goal of the European project, as well as to foster convergence between its different territories, then we must attempt to develop the most efficient and effective policies to fulfill that purpose. The ESIF are, up to now, the main structural tool of the European Union to deliver greater GDP growth in its most economically fragile regions. In order to enhance the outcomes of the cohesion policy and of the implementation of ESIF, we must constantly revise its impacts and the way it is delivering (or not) the desirable results. A negative conclusion with regards to the previous statement would require EU policymakers to rethink and potentially redesign the existent policies and their mechanisms.

As already mentioned – and as will be further described in section 3 -, many authors pursued that effort and strived to achieve some answers for whether the cohesion policy works. However, the effects of the ESIF in the more recent programming periods of their implementation have not received - to the best of our knowledge - the same attention as the previous (earlier) ones. This is notably the case for the last programming period, which run from 2014 until 2020. Furthermore, the evaluation of the impact of ESIF has been extendedly examined in aggregated terms, without accounting for the specificities of different groups of countries and regions, when taking into account their geographical position and, to a large extent, accession date to the EU.

We argue that an overall analysis for the most recent programming period is missing though necessary. Moreover, we believe some path shall be paved as to trace potential different effects of ESIF in distinct groups of countries and regions.

Furthermore, the relevance of undertaking such an analysis is also grounded on the fact that the programming period in question was only the second one in which all the countries from the 2004 and 2007 accessions participated from the very starting point of a program´s implementation. Cyprus, the Czech Republic, Estonia, Hungary, Latvia, Lithuania, Malta, Poland, Slovakia and Slovenia - which officially joined the European Union in May 2004 -, as well as Bulgaria and Romania – which joined the EU in January 2007 -, were all countries subsidized by ESIF during the 2014-2020 programming period (European Commission, 2015). In addition, the 2014-2020 programming period was the very first one that included Croatia, after its accession to the EU in 2013 (European Commission, 2015). All of these countries received EU transfers under the main goal of the ESIF, and most of them had all their regions or almost all eligible to receive EU grants (European Commission, 2015). This observation further strengthens, in our view, the pertinence of the exercise proposed in this thesis. Assessing the impact of the ESIF by accounting for the effects in the regions of the most recent EU Member-States is a work to be further completed.

Having this present, the goal of this Master´s thesis is therefore and firstly to extend the analysis of the impact of ESIF and, by extension, of the EU cohesion policy. In particular, the focus lies on evaluating the effect of the funds delivered to the less developed regions of the EU under the main policy goal of the ESIF during the programming period 2014-2020. In this programming period, the main policy goal was framed as ´Investment for growth and jobs for cohesion policy` (European Commission, 2011). Similarly to the previous programming periods, the biggest slice of funds implemented under the ´Investment for growth and jobs` goal was channeled to the less developed regions of the EU with the aim of promoting economic development and fostering their levels of income per capita (European Commission, 2015). Therefore, the aim of this study is ultimately to answer a question posed in the following terms:

***Have the ESIF implemented under the ´Investment for growth and jobs for cohesion policy` goal promoted economic growth of the less developed regions in the programming period 2014-2020?***

Secondly, this study attempts to shed some light on how the ´Investment for growth and jobs` funds may have had (or not) different effects in less developed regions across separate groups of countries within the EU. Hence, the following sub-research question arises:

*Were there different effects of ´Investment for growth and jobs` funds on the economic growth of less developed regions between distinct groups of regions in 2014-2020?*

In order to answer these questions, this study relies on the mechanisms for casual inference offered by the Regression Discontinuity Design. This design was first introduced by Thistlethwaite and Campbell (1960) and over the years its application in the context of research on economic policies has been expanding (Lee and Lemieux, 2009). As will be outlined in section 4, the very nature of the case under study in this thesis provides a perfect scenario for the use of this methodology.

The main outcome variable of interest of this study is the annual growth of GDP per capita in purchasing power standards (PPS). This variable is not only a suitable measure for the purpose of this thesis, as it is also the most standard indicator used by the literature on the field to assess the impact of ESIF in economic and income per capita growth in less developed regions (Becker et al., 2010; Pellegrini et al., 2013). The effect of interest is that delivered by ESIF invested in less developed regions. The data for this and other variables included in this study is collected from different data sources, the main of those being Eurostat and the regional database of Cambridge Econometrics/Annual Regional Database of the European Commission's Directorate-General for Regional and Urban Policy (ARDECO).

Ultimately, the following hypothesis are to be tested:

*Hypothesis 1*

**The ESIF delivered under the ´Investment for growth and jobs` goal increased GDP per capita growth of less developed regions in the EU.**

With the counterfactual null hypothesis being:

*Hypothesis 2*

**The ESIF delivered under the ´Investment for growth and jobs` goal did not increase GDP per capita growth of less developed regions in the EU.**

And:

*Hypothesis 3*

**The ESIF delivered under the ´Investment for growth and jobs` goal increased GDP per capita growth of less developed regions in the EU in the same way across different groups of regions.**

With the null hypothesis for this being:

*Hypothesis 4*

**The ESIF delivered under the ´Investment for growth and jobs` goal did not increase GDP per capita growth of less developed regions in the EU in the same way across different groups of regions.**

Our findings suggest that there was a positive effect of ´Investment for growth and jobs` funds on GDP per capita (PPS) growth rates of less developed regions in 2014-2020. This effect is estimated to vary between 0,9 and 1,2 percentage points when applying a parametrical approach in the regression discontinuity design, and of 2,5 percentage points with the non-parametric model. No evidence is found for a significant different effect on separate groups of regions.

The remainder of this thesis is organized in the following way: in section 2, a brief historical picture of the EU cohesion policy will be portrayed, and an explanation of its structure, goals, and the different funds will be provided; in section 3, the work already carried out in the field will be revised; after this, the next steps will consist on the explanation of the empirical methodology employed and of the data and datasets used, which is done in sections 4 and 5, respectively; in section 6, the results are displayed; section 7 contains robustness checks of the results; and, finally, sections 8 and 9 will be devoted to further discuss the results and to develop conclusions.

# 2.

# The EU Cohesion Policy explained

## 2.1.
## Historical portrait and reasons behind a cohesion policy in the EU

In the EU – as in other federations and countries - the different actions and mechanisms aiming at fulfilling the goal of promoting greater convergence have historically been referred to under the denomination of ´(EU) regional policy` or ´cohesion policy` (de la Fuente et al., 1995). On the basis of the European efforts towards achieving greater convergence among Member-States and regions (and, therefore, on the need to develop its own cohesion policy), there were mainly political and economic reasons, whose roots can be traced back to the process of European integration and enlargement.

The first time a reference was made to cohesion policies in the European project dates to the very signing of the Treaty of Rome in 1957, which established the European Economic Community (European Commission, n.d.a). In the Treaty´s preamble, the signing parties referred to the will "de renforcer l'unité de leurs économies et d'en assurer le développement harmonieux en réduisant l'écart entre les différentes régions et le retard des moins favorisées" (Treaty establishing the European Economic Community, 1957). In this founding moment of the today´s called European Union, one of the financial instruments nowadays under the umbrella of the ESIF (and specifically related to the cohesion policy) was also launched: the European Social Fund (Federal Ministry of Labour and Social Affairs, 2018), established, in practice, one year later. This fund constituted, however, a small, isolated tool to reach the overall goal of tackling economic and income disparities in the new formed Community (Sutcliffe, 1995). Later, in 1968, the Commission would create for the first time a Directorate- General for Regional Policy, but it was in 1975 that a more significant step was taken towards the institutionalization of a true and broader regional policy in the Community, with the installment of the European Regional Development Fund (ERDF) (Sutcliffe, 1995).

Almost contemporarily to the accession of Denmark, Ireland, and the United Kingdom three years earlier - in 1972-, the ERDF was created with the aim of protecting and creating jobs by delivering investments in "industrial, handicraft or services" activities (Martins & Mawson, 1981: 191). Nevertheless, the new financial instrument was different in shape and goals from the current design of the ERDF, especially due to the use of a system of quotas per Member-State, in which all countries were assigned some part of the ERDF budget (Martins & Mawson, 1981; Bailey & De Propris, 2002), and also due to its main objective, which lied on correcting for the downsides of the different Community actions, rather than attempting to fix regional economic discrepancies (Martins & Mawson, 1981).

After these first landmarks, the European regional policy would again acquire new life around one decade later, with the Single European Act in 1986 and the Reform of the structural funds in 1988. These two events would revolutionize the EU Cohesion Policy and bring it closer to its current shape (Bachtler et al., 2013). Hence, once the Member-States decided to go further on the process of establishing a Single European Market, a revision of the legal basis of the Community started, in which the Single European Act (SEA) came first. For the status of the Community´s regional policy, the SEA was responsible for creating a legal ground at the treaty-level (Sutcliffe, 1995). In the SEA, Member-States agreed upon adding a new Title (Title V) in the Treaty of the European Economic Community, dedicated to economic and social cohesion, with legal precepts highly resembling today´s Article 174 of the Treaty on the Functioning of the European Union (2009). Moreover – and among others – the SEA seems to have consolidated a redefinition of the main goal of the already established ERDF, by explicitly stating its purpose of correcting "regional imbalances (...) through participating in the development and structural adjustment of regions (...) lagging behind" (Single European Act, 1986). The SEA also commanded the Commission to develop a proposal for the reform of the structural funds (Single European Act, 1986).

Following the Single European Act, in 1988 the European regional policy suffered a major reform, especially motivated by, on the one hand, the new accessions of Greece (in 1981), Portugal and Spain (both in 1986) and, on the other hand, the increasing intentions to complete the single market project (Bachtler & Mendez, 2007; Maynou et al., 2016). At this stage, the main political and economic arguments backing the need for a stronger regional policy at the Community level became visible.

Hence, whereas the development and implementation of the European single market was perceived as a force yielding increased economic growth, it also elicited doubts about the possible negative impacts on the weaker economies of the EU (Rodriguez-Pose and Fratesi, 2004). As Bailey and De Propris (2002a) explain, there was a strong need to support the less developed regions across the EU territory and, in order to move forward with the completion of the single market, it was crucial to compensate those regions affected by the imbalances arising from greater integration. By the time the single market was being planned - and later introduced -, the response to these fears was delivered in the form of a range of programs and financial instruments designed to support the less developed countries and regions, more likely to be negatively affected by that new giant step in the process of the European economic integration (de la Fuente et al., 1995). Moreover, a second line of economic reasoning voiced out that the development of the weaker economies in the new European single market would also be beneficial for the economic activity of the community as a whole, including those economies of the wealthiest countries (Bailey and De Propris, 2002a).

In addition to these pure economic ideas, advancing with the single market also required political support from the different Member-States. With this view, it was necessary to guarantee that the new project for a single market was perceived as fair by the involved parties and, again, the development of more ambitious regional policies emerged as response to foster political endorsement for deeper integration (De la Fuente et al., 1995; Bailey and De Propris, 2002a).

Concretely, the reform of 1988 unified all the distinct structural funds devoted to cohesion purposes under a common EU cohesion policy (European Commission, n.d.a). It also doubled the available budget for the new unified cohesion policy (Sutcliffe, 1995; Bachtler & Mendez, 2007), and introduced a new framework of governing principles of the cohesion policy, namely: a shift towards multi-annual planning of the funds spending; a greater concentration on the less developed regions; a more strategic application of funds, in which added value should arise from the investments (the so-called additionality principle); and strengthen cooperation between the supranational, national and local levels on the implementation of the funds (Bachtler & Mendez, 2007; European Commission, n.d.a).

After the 1988 reform and until today, the relevance of the cohesion policy never ceased to increase. The Maastricht Treaty of 1993 created a new financial instrument to be added on the existing group of ESIF allocated to the cohesion policy: the Cohesion Fund (Boldrin and Canova, 2001). In the programming period 1994-1999, the available funding was doubled and reached a third of the total EU budget (European Commission, n.d.a). In 2007-2013, the expenditures with the cohesion policy became the biggest slice of the EU budget (Benedetto, 2019). Today, a record value of 392 billion euros is allocated to the EU cohesion policy, most of them delivered through the ESIF (European Commission, n.d.).

## 2.2.
## The functioning of the cohesion policy and the programming period 2014-2020

The functioning of the cohesion policy is complex. It is also intrinsically related to the planning and structure of the main financial instruments that support it: the European Structural and Investment Funds.

Nowadays, there are essentially four financial instruments that support the cohesion policy, all part of the broader group of ´European Structural and Investment Funds`. These four funds are: the European Regional Development Fund (ERDF), the Cohesion Fund (CF), the European Social Fund Plus (ESF+), and the Just Transition Fund (JTF) (European Commission, n.d.). This setup was not always constant. Nevertheless, since the regional policy reform of 1988 that it hasn´t changed substantially (Bailey and De Propris, 2002a). As mentioned in the previous section, after this reform, the ESIF devoted to the cohesion objectives were unified under a unique EU cohesion policy (European Commission, n.d.a). When the reform took place, the ERDF and ESF already existed; later, in 1993, the Cohesion Fund was introduced (Boldrin and Canova, 2001; European Commission, n.d.a).

These available funds for the EU cohesion policy are (as they were in the past) attached to the EU budget (Benedetto, 2019). Once the Single European Act entered into force in 1987, the structure of the EU budget changed substantially and with it the functioning of the cohesion policy. From that point onwards, the EU budget started to be planned on a long-term basis, by means of new multiannual programming periods that were introduced and which incorporated time windows of several years (Benedetto, 2019). With the Treaty of Lisbon, these programming periods were renamed to the today´s called ´Multiannual Financial Frameworks` (Benedetto, 2019).

After the introduction of these long-term budgets - and the first one started to run in 1988 -, the cohesion policy and the management of the different ESIF allocated for cohesion purposes would also begin to be planned in ´programming periods`, each of them falling within the different EU multiannual budgets (Hagen and Mohl, 2009). The first programming period ran from 1989 to 1993 and, after that, other four programming periods followed and have already been completed: 1994-1999, 2000-2006, 2007-2013, 2014-2020. Currently, a new programming period is underway, having started in 2021, and will last until 2027 (European Commission, n.d.b).

In each programming period, the sums from the different ESIF available for the cohesion policy are allocated on the basis of a European Nomenclature of Statistical Territorial Units. Using this Nomenclature, territories eligible for receiving funds are defined (Boldrin and Canova, 2001). Level 2 of the Nomenclature (NUTS 2) has historically been the most relevant for the cohesion policy, as it is the one used by the EU to allocate the biggest slice of the funds devoted to the cohesion policy. This major amount of funds is channeled for regions across the EU – defined precisely at a NUTS 2 level. Other amounts of the available funds are, for instance, allocated at the Member-State level.

Moreover, the most significant part of the available funds is not only delivered to regions across the EU (as explained, using the NUTS 2 level), as it is also allocated under a specific policy goal that has historically been the most important when planning the cohesion policy and the implementation of ESIF. Thus, when the funds are applied, they not only derive from different sources (today, the four funds referred above), as they also aim at accomplishing several policy goals, which are defined previously to the implementation of each programming period. Over the years, one of those policy goals has particularly attained special relevance, and it has also been shown to be the one under which the biggest slice of funds is allocated.

In the past, this policy goal was referred to as the ´Objective 1` of the cohesion policy and the ESIF. In the programming period 2007-2013, this objective was reframed to the ´Convergence` goal. In the programming period 2014-2020, it was renamed to ´Investment in Growth and Jobs for cohesion policy` (European

Commission, 2015). In spite of its different formulations, the main feature of this policy goal is that it essentially aims at supporting the less developed regions of the EU in economical and development terms (European Commission, 2007).

In the context of this main policy goal, an important aspect to be considered is that of how the EU has been determining which regions are to be considered as less developed and therefore eligible for receiving the biggest part of available funds. The criterion has remained constant since this policy goal was launched under the formulation of ´Objective 1`: having a GDP per capita below 75% of the EU average.

Furthermore, it is to be noted that the categorization of regions as less or more developed accordingly to the ´75% rule` is usually carried out at some point in time preceding the implementation of a given programming period.

*Figure 1* captures this overall architecture of the cohesion policy and the ESIF. Moreover, it describes how this scheme specifically applied for the programming period to be analyzed in this study: 2014-2020.

As the figure portrays, in the programming period 2014-2020, the main policy goal was framed as the ´Investment for growth and jobs for cohesion policy` (European Commission, 2011). The criterion used to classify regions in one of the three development categories remained exactly the same as in the previous programs. Hence, the less developed regions were those with a GDP per capita below 75% of the EU average. In turn, regions with levels of GDP per capita above that threshold would either be transition or more developed territories (European Commission, 2011). In the programming period 2014-2020, the application of this criterion and the eligibility to be awarded with funds under the ´Investment for growth and jobs goal` was defined in the years 2007-2009 (Regulation 1303/2013, 2013).

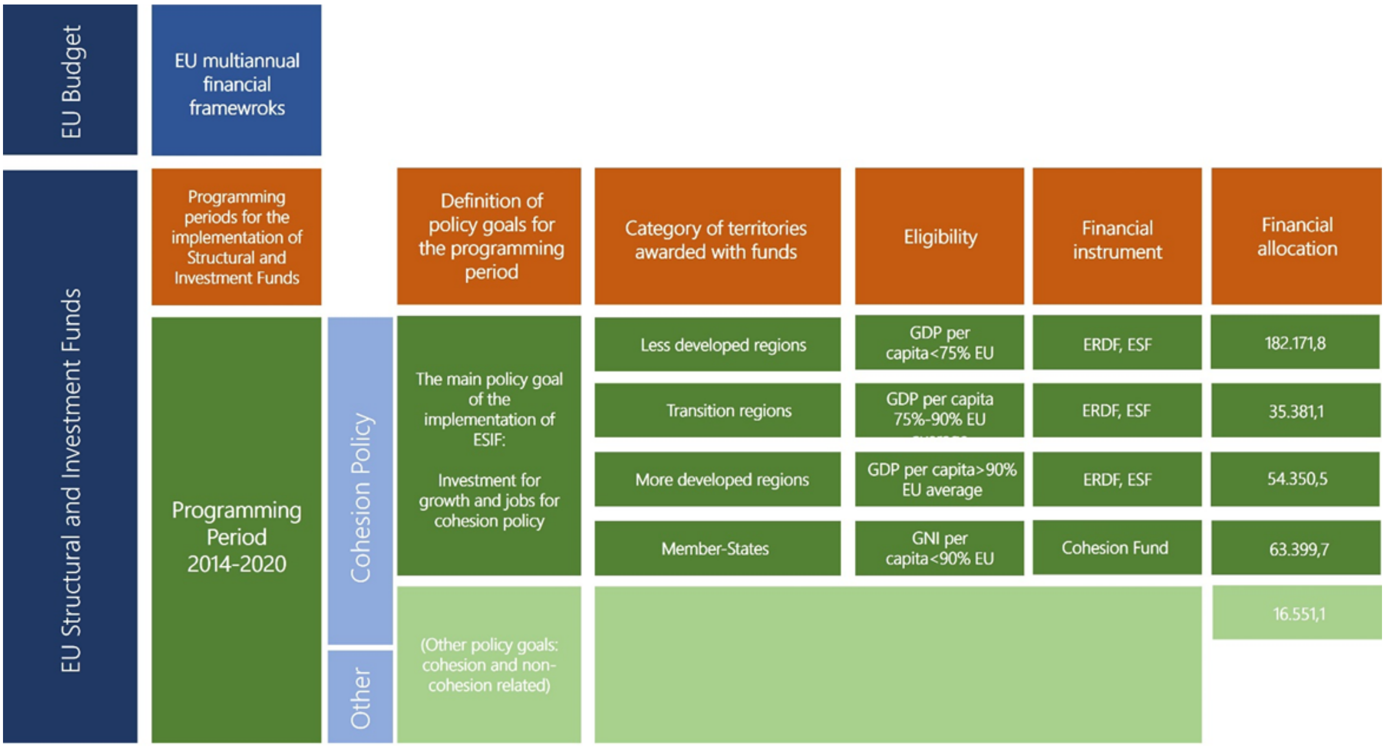
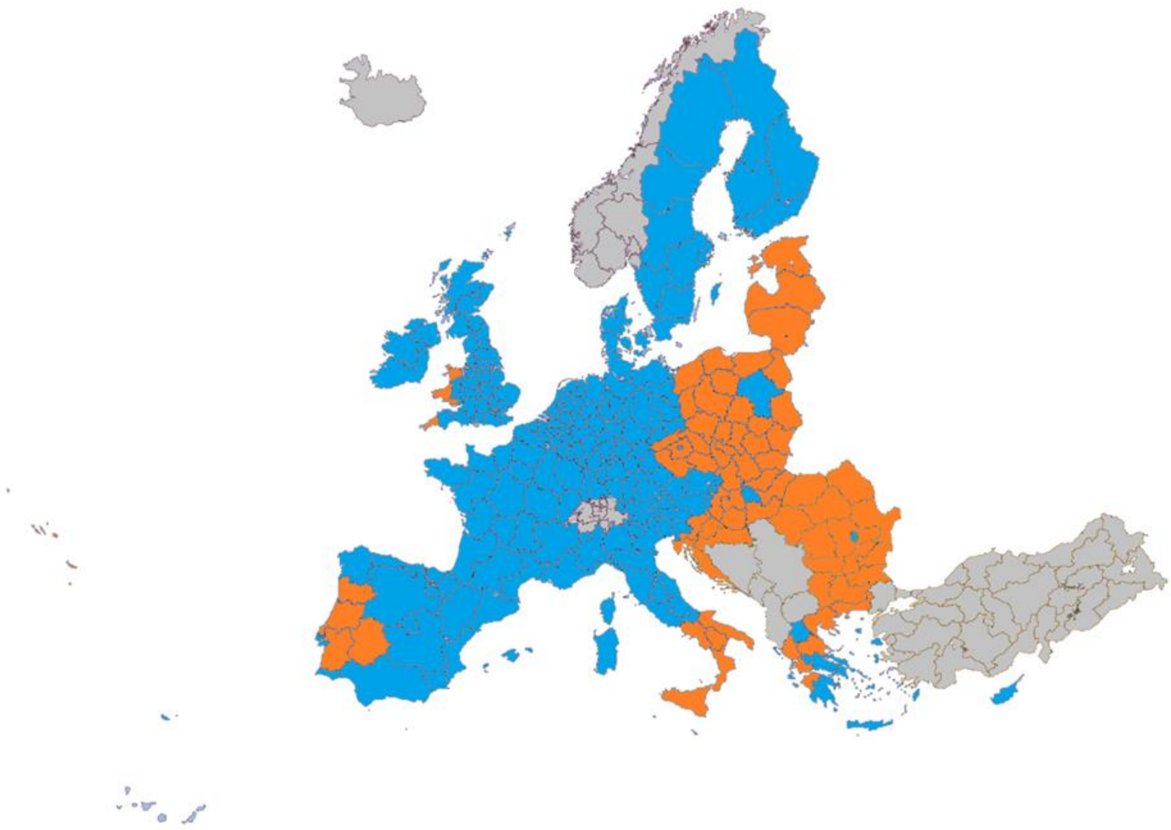



*Figure 1. The functioning of the cohesion policy and the ESIF, and the programming period 2014-2020.* The figure depicts how the European Structural and Investment Funds are implemented and the relationship of these funds with the cohesion policy. It also portrays how the general architecture of the ESIF available for cohesion policy applied to the programming period 2014-2020. In orange, the information on the structure and functioning of ESIF. In green, how that structure specifically applied to the programming period 2014-2020. The amounts are given in Million Euros, current prices.

In line with the historical mission of the funds applied under the main policy goal – in 2014-2020, the ´Investment for growth and jobs for cohesion policy` - the biggest slice of available funds was devoted to less developed regions. These regions were awarded more than two times the amounts allocated to transition and more developed regions taken together (European Commission, 2015). This calls for the observation that less developed regions were heavily more treated by ESIF compared to the other territories.

*Map 1* depicts the treated and untreated regions during the programming period 2014-2020. The treated regions are those in orange, which consist of less developed regions, awarded with a significantly greater amount of ESIF – as shown in Figure 1 . In turn, the untreated regions – in blue – consist of both transition and



***Map 1.*** *The treated and untreated regions in the programming period 2014-2020.* The map shows Europe´s NUTS 2 regions accordingly to the nomenclature in force in 2006, just before eligibility for the funds in the programming period 2014-2020 was defined. The map distinguishes between treated and untreated regions during the programming period 2014-2020 of the cohesion policy and the ESIF. The first are represented in orange; the second, in blue.

more developed regions, which even though were also awarded with EU funds, this value was drastically lower than that allocated to less developed territories.

As will be described in section 4, the functioning of the cohesion policy and, in particular, the features of the allocation of ´Investment for Growth and Jobs` funds, give rise to a scenario in which the effect of those funds on the economic growth of less developed regions can be assessed by means of a Regression Discontinuity Design.

# 3.

# Literature Review

The attempts to assess the impact of the EU´s cohesion policy and the ESIF were already multiple heretofore. In published work on the subject, there is a divide between studies that tried to assess to which extent convergence has occurred in the European Union, and those which focused on evaluating the impacts of ESIF on economic growth (without necessarily accounting for the existence, or not, of convergence). Nevertheless, for both approaches, a consensus is hard to establish as to whether the cohesion policy is delivering the desired outcomes. Indeed, the empirical findings so far reached have suggested different conclusions, which is often attributed to the variety of models, data and econometric approaches applied (Esposti and Bussoletti, 2008; cf. Hagen and Mohl, 2009; Crescenzi and Giua, 2020). Hence, some studies pointed to positive significant effects of cohesion policy and ESIF, others to nonsignificant effects, and some even to negative effects of the ESIF. Moreover, whereas part of the available studies approached the overall impact of the EU cohesion policy, others specifically addressed the effect of receiving funds under the main policy goal of the ESIF.

The work of Sala-i-Martin (1996) is underlinable has a referential starting point in the chain of existing studies on the subject. In his study on regional economic convergence in several countries and the EU, Sala-i-Martin (1996) found an average convergence rate similar to other countries in the globe (almost 2%) for the group of EU regions that was studied. In turn, exploring the impact of the EU cohesion policy in the less developed regions, and by accounting for aggregated EU transfers provided under different funds and objectives, Boldrin and Canova (2001) found no evidence that the ESIF have contributed for income per capita convergence but, instead, that long-run growth rates of both less advantaged and more developed EU regions are expected to remain similar. Focusing on a ten-year interval covering the first two programming periods (from 1989 to 1999), Dall'erba and Le Gallo (2008) reached similar conclusions as to the influence of ESIF. The authors found that convergence took place among EU regions, but that the ESIF did not offer any significant contribution in that direction.

Focusing solely on the main policy goal of the cohesion policy and the ESIF, Becker et al. (2010) assessed the impact of the funds delivered under the former Objective 1 (the equivalent to the today´s ´Investment for growth and jobs` goal). The authors relied on a fuzzy Regression Discontinuity Design to evaluate the effects of Objective 1 funds which - we remind - represent the main channel of support for lagging regions; they performed this assessment for three programming periods (running from 1989 to 2006) by using a parametrical approach. The empirical results of Becker et al. (2010) suggest that Objective 1 funds did not impact employment growth but did have a statistically significant positive effect on GDP per capita growth of the regions that received the funds. In particular, the findings of Beckert et al. (2010) suggest that Objective 1 funds increased the average growth of GDP per capita of treated regions in around 1,6 percentage points.

Pellegrini et al. (2013) did also implement a Regression Discontinuity Design to assess the impact of Objective 1 funds on GDP per capita growth of the less developed regions from 1996 to 2006. In opposition to Becker et al. (2010), however, these authors applied a non-parametrical model as the main methodological strategy. The results reached by Pellegrini et al. (2013) suggest that the ESIF delivered a 0,6 percentage points yearly higher economic growth for less developed regions. Using the same methodology, Crescenzi and Giua (2020) found similar results (0,36) for the period 2000 to 2010.

Esposti and Bussoletti (2008) performed the same analysis on the effects of Objective 1 funds, but by means of a panel-data approach. The authors found a positive – though short – impact of ESIF in Objective 1 regions. With the same focus but opposite results, the findings of Breidenbach et al. (2019) suggest a negative return rate between 0 and -0,5% of ESIF investments in lagging regions in the years 1997 to 2007.

Whilst the above studies focused on evaluating the overall impact of the ESIF (either on global convergence or solely on the economic growth of less developed regions), other authors attempted to explain the specific outcomes on certain policy fields. This was the case of Ferrara et al. (2017), who exploited the effects of the EU cohesion policy on research and innovation, as well as on transport accessibility, during the programming period 2000-2006, and simultaneously accounting for lagged effects on the years following that timespan. The authors relied on a regression discontinuity design and found a statically significant positive effect of EU transfers on both research and innovation activities and transport accessibility in Objective 1 regions.

Furthermore, a group of scholars narrowed down the scope of assessment to within-country or sub-country effects. Crescenzi and Giua (2020) evaluated the specific country-effect of funds delivered to less developed regions between 2000 and 2010. Their empirical findings suggest that ESIF promoted income growth in German regions (3,5 percentage points) but not in Italian, Spanish and British regions (where no significant effect was found).

Studying the impact of the ERDF in income per capita levels in Spain, De la Fuente et al. (1995) concluded that between 1986 (when the official accession of Spain to the ECC took place) and 1990, the European Regional Development Fund delivered an increase of around 0.34 to 1.40 percentage points in regional income per capita, when compared to the hypothetical scenario in which the country would have not benefited from EU transfers. Lolos (2009) described a significant positive effect of ESIF in regional growth in Greece from 1990 to 2005. More recently, Aiello and Pupo (2011) evaluated the effect of ESIF in Italy in the period 1996-2007. Their empirical findings showed a minor but positive effect of ESIF on regional convergence in Italy, with greater impact in terms of GDP per capita growth in the south region (the main beneficiary of EU funds in the country).

With a double narrower scope, Barone et al. (2016) focused solely on the Italian region of Abruzzi and on assessing the impact of leaving Objective 1 status. The authors compared the region of Abruzzi (which in 1997 completely ceased to receive Objective 1 funds, after having been eligible to - and effectively received - EU transfers from 1989 to 1996) to the group of the other southern Italian regions (the Mezzogiorno regions) which continued to receive EU grants under the Objective 1 of cohesion policy. In their study, the authors found that after having ceased to receive Objective 1 funds, the region of Abruzzi had lower levels of GDP per capita growth when compared to other southern regions which remained under Objective 1 treatment.

An outstanding innovative contribution was that of Rodriguez-Pose and Fratesi (2004) which, focusing on Objective 1 regions, exploited not only the overall effect of ESIF in less developed EU regions, as also approached the specific impacts of different investment options. In particular, Rodriguez-Pose and Fratesi (2004) researched on the effects of EU transfers in Objective 1 regions of investments undertaken under four priority axes: "support to agriculture and rural promotion", "business and tourism support", "investment in education, re-qualification and all measures targeting the human capital of the region" and "investment in infrastructure, transport, and environment" (Rodriguez-Pose and Fratesi, 2004: 105). Their results show that investments in agriculture and rural promotion led to a positive impact in the short-term, as well as investments in human capital/education, which also had slight positive significant effects on the medium-term, whereas the transfers allocated to infrastructures/transports/environment and business/tourism (the two main axes of investment) were not reflected on any significant consequence for the GDP per capita growth of Objective 1 regions. The approach carried out by Rodriguez-Pose and Fratesi (2004) is also highlightable as it called upon attention to the fact that the application of ESIF might generate results which are only observable in the medium and long-term. For this reason, the Authors considered the interval of years from 1989 to 1999 (corresponding to the first two programming periods) and proceeded to a cross-sectional analysis for different timespans within that interval, and to a panel data analysis with annual lags up to seven years (leading to the results previously described).

Other studies further explored what Dall`erba and Le Gallo (2008) eloquently named the "conditional effectiveness" of the funds, one of those also cited by these Authors: the work of Ederveen et al. (2006). A first note to be given on the study of Ederveen et al. (2006) is that the scope of analysis was set at the country level, instead of assessing regional growth. As the authors explain, among other virtues, this option potentially reduces the bias generated by spillover effects. As for the contribution on evaluating the impact of EU grants, Ederveen et al. (2006) found empirical evidence backing the idea that the effectiveness of ESIF depends on the quality of the institutions of a Member-State: for those countries performing better on institutional quality standards, the ESIF delivered greater economic growth.

From a different angle, Becker et al. (2018) addressed the effectiveness of the ESIF during periods of financial and economic crisis. The Authors applied a fuzzy RDD to assess the effects of Objective 1 treatment during the period 1989-2013 (comprising four programming periods) and found a positive significant effect of a bit less than a 2 percentage points increase of GDP per capita growth. In a second step, the authors isolated the subperiod 2000-2013 (which included the two programming periods 2000- 2006 and 2007-2013) and added government-bond-yield spreads (GBYS) as a variable to reflect the distinct effects of the 2008 financial and economic crisis in separate regions. The authors found that without including the new variable, the results for GDP per capita growth in the period 2000-2013 were smaller than those for 1989-2013. When accounting for the different intensities by which countries were hit by the crisis from 2008 onwards, the authors empirically discovered that the Objective 1 treatment effect on GDP per capita growth became lower but increased for the outcome variable employment growth. As referred by the authors, Objective 1 treatment effect was fully canceled by "an increase in GBYS by about one-and-a-half standard deviations" (Becker et al., 2018: 148).

The contribution of Becker et al. (2018) is also noteworthy for their original assessment of the effect of ´entering` or ´losing` Objective 1 treatment in income per capita growth, employment growth and other economic variables. As for the evaluation of the impact of ´entering Objective 1` status, the authors compared those regions which became eligible and started to receive EU transfers at a certain point in time between 1989 and 2013, to those which never became eligible; in turn, for the effect of ´losing Objective 1` treatment, the authors compared the regions which ceased to be eligible for Objective 1 funds to those which remained under the eligibility criteria during the period under analysis. Their empirical findings revealed that regions ´entering` Objective 1 at a given point exhibited a GDP per capita growth of 2.1 to 2.6 percentage points higher than "never-treated" regions, whereas those which lost the access to Objective 1 funds had a GDP per capita growth 1.7 percentage points lower than "always-treated" regions (Becker et al., 2018: 150).

Lastly, one mention for the works of Breidenbach et al. (2019), Becker et al. (2010), and Ederveen et al. (2006), for the way they considered the potential bias generated by spatial spillover effects. These effects relate to the idea that the investments made in less developed regions may not only benefit themselves but also other territories in the proximities. The consequence of ignoring this fact is that the results may be downward biased, as the treatment delivered by the funds is also felt by many other regions, making the estimation of the real effect more imprecise. As previously mentioned, Ederveen et al. (2006) tried to overcome this barrier by focusing on the country level. Becker et al. (2010) included a robustness check in their results controlling for potential spillover effects. They did this by two different means: first, by excluding non-less developed regions within a radius of 150-200km next to less developed regions that were awarded funds; second, by means of an indicator variable for regions with neighboring funded regions. Their results showed no significant effect of controlling for spatial spillovers in the final estimations of the impact of ESIF in less developed regions. Breidenbach et al. (2019) find that spatial spillovers do occur and have a negative effect on the estimation of the impact of ESIF (which, as described before, was found to be between 0 and -0,5% by these authors).

# 4.
# Methodology

To assess the impact of the ESIF delivered under the ´Investment for growth and jobs` goal on income per capita levels in less developed regions in the EU, we rely on the mechanisms for causal inference offered by the regression discontinuity design (hereinafter also: RDD or RD design), introduced by Thistlethwaite and Campbell (1960).

The essence of a RDD lies on the idea that, although real life does not always provide a scenario in which comparisons between perfectly randomized treatment and control groups are possible, there are cases in which the assignment of units into each of those groups can be considered almost as good as random (Lee and Lemieux, 2009). These are cases in which a cut-off point exists for a given variable (the ´assignment variable`), and in which falling under or above that cut-off point is the only condition determining whether a unit receives or not a specific treatment (Lee and Lemieux, 2009; Angrist and Pischke, 2014). The criteria established by the EU to assign the biggest slice of ESIF fits in the range of cases in which a cut-off point exists and determines the composition of comparable treatment and control groups.

To recap - and as explained in section 2.2 -, the EU uses a ´75% rule` to determine which regions are considered as less developed, as opposed to transition and more developed ones: those with a GDP per capita below 75% of the EU average are less developed, whereas those above that value are either transition or more developed regions. The regions classified as less developed for each programming period are allocated with a significantly bigger slice of funds. In the programming period 2014-2020, the EU also used the ´75% rule` to determine which regions were considered as less developed for that period. This categorization was performed based on each region´s average GDP per capita (PPS) in the years 2007-

2009, which was when eligibility for receiving funds in the period 2014-2020 was defined. During the implementation of the 2014-2020 programming period, the biggest slice of funds was then applied in those regions.

Within this picture, the application of the ´75% rule` gives rise to a scenario in which the condition of being treated or untreated only depends on whether a given region falls below or above the ´75%` threshold. If a region had a GDP per capita below 75% of the EU average when eligibility was defined, it received treatment during the implementation of the 2014-2020 programming period (i.e., that region was hardly financed by ESIF). In opposition, if a region had a GDP per capita above that value, it was not treated (it did not receive a significant amount of ESIF). The variable that determined whether a region was awarded funds or not was the region´s initial GDP per capita, i.e., the region´s GDP per capita when eligibility was determined before the programming period. The ´75%` threshold is therefore the cut-off point of our case, whereas a region´s initial level of GDP per capita in PPS is the ´assignment variable`, in which receiving treatment or not exclusively depended.

In this context, we can identify treatment and control groups for assessing the impact of ESIF on lagged regions in the 2014-2020 programming period. Hence, the treatment group is composed of the less developed regions themselves, which received the treatment. In turn, the control group is integrated by the transition and more developed regions, which were untreated regions. The assignment of each region into one of the relevant groups for the comparison exclusively relied on in which side of the ´75%` threshold a region fell given its initial level of GDP per capita (PPS).

Furthermore, in a RD design, the focus is in the discontinuity that is found at the cut-off point between the treatment and the control groups. When this discontinuity exists, there is ground for the application of a RDD, as the jump observed at the cut-off point can be interpreted as the effect of the treatment on the treated (for instance, the impact of an economic policy) (Lee and Lemieux, 2009;

Pellegrini et al., 2013; Ferrara et al., 2017). The challenge for the researcher is therefore to estimate the ´size of the jump` at the cut-off point and to verify whether it is really due to the treatment or, instead, to any other factor with an impact on the outcome.

In addition to these considerations, in a RDD we are particularly interested in looking to those observations which fall close to the relevant cut-off. By narrowing down the number of observations to those units which are closer to the threshold value, one will probably find a set of units which are very similar in all aspects, but which differ on their treatment status. In this context, a more ´apples to apples` comparison is reached, and the difference in the observed outcome between the units on both sides of the threshold can be attributed solely to whether a unit was treated or not.

Bringing this observation to our case, it is likely that those regions with very similar levels of GDP per capita (PPS), although falling in different sides of the cut-off when the ´75%` is applied, they probably share the same profile in what regard a range of characteristics that could potentially affect the relevant outcome. Illustratively, a region with a GDP per capita (PPS) of 74% probably resembles another one with a GDP per capita (PPS) of 76% in a series of economic, social, demographic, geographical conditions, but diverges in one aspect: the first is eligible for ESIF under the condition of "less developed region" - and is therefore awarded with the treatment -, whereas the second is not. Since these regions are very similar in all other factors that could potentially influence the outcome, the distinct growth rates of GDP per capita (PPS) between them can be explained by whether the region received, or not, a substantial amount of funds (i.e., if the region was treated or not). The difference in the outcomes will be – again - the (treatment) effect of those funds.

Having said this, the model followed in this study is a sharp RD design, as the condition of being treated exclusively relies on the assignment variable ´level of initial GDP per capita (PPS)` and in which side of the cut-off point a region fell, accordingly to the assignment variable (Hahn et al., 2001; Imbens and Lemieux, 2007). Following the approaches of Pellegrini et al. (2013) and Ferrara et al. (2017), we make sure this assumption holds by excluding the very few observations that (for reasons that were not clearly identified) received any form of treatment even though they did not comply with the ´75% rule`, and vice-versa. In order to estimate the treatment effect, we apply a parametrical approach in the regression analysis and a non-parametrical approach as a robustness check of our results.

In the context of a RD design, a parametrical approach makes use of all the observations of the case and tries to estimate the treatment effect by looking for the optimal functional form that captures the relationship between the outcome and the assignment variable (van der Klaauw, 2008; Jacob et al., 2012; Cattaneo et al., 2019). The advantage of this model is that by implementing flexible non-linear specifications, it excludes the possibility that the discontinuity at the cut-off point is due to undetected non-linearities of the relationship between the outcome and assignment variable (Pellegrini et al., 2013).

The most important challenge for RD practitioners when implementing parametrical models is to choose the appropriate functional form to be employed, i.e., to decide on the most suitable polynomial order of the regression (Lee and Lemieux, 2009; Pellegrini et al., 2013). Usually, several estimates are presented for regressions of different polynomial orders (Becker et al., 2010). Moreover, some tests can be performed to assess the goodness of fit of the functional form employed, including informal strategies and cross-validation procedures (Lee and Lemieux, 2009; Jacob et al., 2012).

Considering this, we want to estimate the value of the parameter $\beta$ in $Treat_{it}$ (treatment effect) in a regression of the following kind:

$$Growth_{it} = \alpha + \beta Treat_i + f^m(\beta AsVar_i) + \lambda_i + \sigma_t + \beta Cont_{it} + \varepsilon_{it}$$

Where $Growth_{it}$ stands for the dependent variables of interest, i.e., annual growth of regional per capita GDP in PPS of region i in year t or employment growth of region i in year t. $Treat_{it}$ is a treatment dummy, taking the value one (1) when a region is treated (i.e., hardly financed by ESIF), and zero (0) otherwise. $AsVar_{it}$ is the initial level of GDP per capita in PPS of a given region before the starting of the programming period, which is the assignment variable of the case. The relationship between the dependent and the assignment variable is captured by a function $f$ of $m$ polynomial order. $\lambda_i$ and $\sigma_t$ are country and year fixed effects, respectively. $Cont_{it}$ stands for additional control variables included in the model. $\varepsilon_{it}$ is an error term, which is assumed to be uncorrelated with the treatment variable. Once more, the parameter $\beta \text{ in } Treat_{it}$ is the relevant treatment effect to be estimated.

The ´Akaike information criterion` (AIC) is used as a test of the goodness of fit of the regressions of different polynomial orders applied (Lee and Lemieux, 2009; Jacob et al., 2012).

In addition to performing the parametric approach in a standard way, this study also introduces an element of novelty in the analysis, by attempting to estimate whether the ´Investment for growth and jobs` funds deliver different results between less developed regions across distinct groups of countries. This is achieved by interacting the treatment dummy variable with additionally created categorical variables standing for groups of countries on a basis of their geographical position (and largely also entrance date in the European Union).

Specifically, three groups of countries are considered: the southern/Mediterranean countries (Portugal, Spain, Italy, Greece, Cyprus and Malta); the eastern countries and Balkans (Bulgaria, Czech Republic, Latvia, Lithuania, Hungary, Poland, Romania, Slovenia, Slovakia, Estonia and Croatia); and the central and northern countries (Belgium, Denmark, Germany, Ireland, France, Luxembourg, The Netherlands, Austria, Finland, Sweden and the United Kingdom). This last group is used as a reference for the regression analysis with interaction effects.

The main goal of extending the analysis by incorporating these interaction effects is to assess whether EU funds might have delivered (or not) different effects across different types of regions (on the basis of the assumption that their geographical position, as well as other cultural and economic reasons might also be relevant for the effectiveness of the implementation of the funds).

Subsequently, the non-parametric approach is applied. In contrast with the parametric model – which relies on using all the available observations – the non-parametric approach is focused solely on analyzing those units close to the cut-off. The underlying idea in the non-parametric approach is that for those observations closely around the threshold there is no significant concern regarding the optimal functional form that captures the relationship between the outcome and the assignment variable. Instead, in the narrower windows of units around the cut-off, it is very likely that the proper functional form is close to linear (Jacob et al., 2012). In this context, estimating the treatment effect can be done by running simple linear regressions (Jacob et al., 2012; Cattaneo et al., 2019). The treatment effect is given by the difference of estimates for the observations on each side of the cut-off (Lee and Lemieux, 2009).

Furthermore, in the non-parametric approach two questions acquire particular relevance: first, the kernel function applied in the regression; second, the choice of the bandwidth size.

Regarding the kernel function employed, we follow the reference literature on the field and employ different kernels as to compare the reached results (Pellegrini et al., 2013; Ferrara et al., 2017; Cattaneo et al., 2019). Specifically, rectangular, triangular and epanechnikov kernels are used.

As mentioned before, the ´bandwidth` stands for the range of observations around the cut-off that are included in the analysis (Angrist and Pischke, 2014). A trade-off between validity and the necessary variation is here to be underlined. Hence, on the one hand, selecting a range of fewer observations falling closer to the cut-off value leads to increased reliability of the conclusions to be drawn, as the considered observations reveal a higher probability of being similar in all aspects except their treatment status. On the other hand, however, a shrink in the number of observations has an effect on the levels of variation one needs to perform the regression analysis and, if disproportional, might lead to less precise or insignificant results (Angrist and Pischke, 2014).

When performing the non-parametric approach and in line with the nature of the RD design, several bandwidths are implemented to further check the robustness of the results. Thus, first, an optimal bandwidth is computed using the method proposed by Imbens and Kalyanaraman (2012) to reach optimal data-driven bandwidths. In addition, other ad hoc windows of observations are also selected.

# 5.

# Data

## 5.1.
## Databases and collection of data

Data used in this study is mainly derived from Eurostat, the official statistical office of the European Union. As will be described, data for the outcome variables and for other covariates included in the robustness checks is derived from this database. In addition, the regional database of Cambridge Econometrics/Annual Regional Database of the European Commission's Directorate General for Regional and Urban Policy (ARDECO) is used as a complementary source for missing observations in both the outcome variables and covariates. Moreover, information regarding the categorization of a region´s development level and whether it was treated or not by ESIF is collected from official documents and reports of the European Commission.

The data was mainly collected for the central level of interest of this study: the regional level. Data at the country level was also collected for descriptive purposes (developed in section 5.4), as well as for computing the average EU GDP per capita (needed to control for compliance with the ´75% rule`).

The relevant regional level for this study is the NUTS 2 typology of the European ´Nomenclature of territorial units for statistics`, at which the EU allocates the big majority of ESIF and defines regional eligibility to receive EU transfers, accordingly with a classification between less developed, transition and more developed territories (in which the first – we remind – are the main beneficiaries). As referred in section 2.2, for the programming period 2014-2020, regions were classified under one of these three categories based on their economic outcomes of 2007-2009 (Regulation 1303/2013, 2013). The data was collected for all NUTS 2 regions of the EU 28 (in 2013) Member-States.

Nevertheless, the collection of data for NUTS 2 regions was not entirely straightforward. Hence, by the time this study was developed, the regional data available on Eurostat was (almost completely) offered accordingly to the Nomenclature of territorial units for statistics in force in 2021, which since 2009 had been revised in four moments (2010, 2013, 2016 and 2021). Throughout these revision processes, geographical modifications occurred, with some regional boundaries redefined. Consequently, the available data was presented for a European NUTS 2 map which did not entirely correspond to that of when eligibility was defined and according to which funds were implemented. Using the available data without accounting for the geographical modifications that took place would be problematic for this study. Firstly, some regions would no longer comply with the ´75% rule` after the geographical redefinition. Secondly and due to this, the treatment effect would be biased. It was therefore necessary to account for geographical modifications and their impact on the data.

## 5.2.
## Outcome, explanatory and assignment variables

### a) Annual growth of regional GDP per capita in purchasing power standards

The main outcome variable of interest is annual growth of the gross domestic product per capita in purchasing power standards (GDP per capita in PPS). In addition to the centrality of GDP per capita in the EU cohesion policy, its annual growth rates are also a common indicator – the most frequent, to the best of our knowledge - used by the literature on the field to assess the impact of EU funds in income levels. Some studies rather use average annual growth of the same indicator, although with minor differences with regards to the achieved results (cf. Becker et al., 2010; Pellegrini et al., 2013).

The gross domestic product (GDP) is the most used measure of the economic activity, generally standing for the value of goods and services produced in a territory (Stiglitz et al., 2009; Lequiller and Blades, 2014). The per capita value of GDP (given by the total GDP divided by the corresponding population) provides an indication of the "living standards or economic well-being" (Organization for Economic Co-operation and Development, 2009) and is an indirect measure of income per capita (The World Bank, n.d.). GDP growth is the standard measure of changes in economic activity (OECD, 2009); when computed at a per inhabitant level, it offers an indication as to how living standards of individuals evolved.

The data on GDP per capita was collected from Eurostat for the years 2013 to 2020. Missing data for the whole period for the United Kingdom´s NUTS 2 regions, as well as for the years of 2013 and 2014 in the case of France, is derived from the regional database of Cambridge Econometrics/ARDECO. The annual growth rate of GDP per capita in PPS was next computed for the regions and the years analyzed.

b) Annual growth of regional employment

Annual regional employment growth is added as a secondary outcome variable. We follow other studies on the field in adding this variable as an alternative outcome (cf. Becker et al., 2010; Becker et al., 2018).

This variable stands for the percentual annual change of total individuals employed with fifteen and over fifteen years-old in each region (with numbers given in thousand persons). Equally to the GDP, total employment is a main indicator of the state of an economy. Computing its annual variation offers an indication as to how the economy developed between years.

The data for this variable is derived from Eurostat for the years 2013-2020, with the exception made for data for all regions of the United Kingdom in the year 2020, in which case data is collected from the regional database of Cambridge Econometrics/ARDECO.

c) **Treatment dummy variable– Hardly and Non-hardly financed regions**

For the sake of clarity, this is the point to remind that the purpose of this study is to assess the impact of EU funds delivered to less developed regions under the ´Investment for growth and jobs` goal. These funds are therefore the explanatory variable we are interested in. In other words, we can refer to them as the policy ´treatment` whose effects are to be assessed.

During the programming period 2014-2020 and under the ´Investment for growth and jobs` goal, that ´treatment` was delivered to some regions (the less developed), but not to others (the more developed ones). The regions that received the funds – i.e., the treatment – are said to be ´treated regions`, whereas those which were not granted with EU transfers – i.e., did not receive the treatment – are designated as ´untreated regions`.

In the present study – and in line with the RD design -, this distinction between ´treated` and ´untreated` regions is expressed in the form of a dummy variable, in which treated regions are assigned with the value "1" (one), and untreated regions with a "0" (zero) instead. The value assigned to a given region is therefore an expression of whether the treatment was implemented in that region or not. If a region is assigned with the value 1 (one), it was subject to the treatment, i.e., the region received EU funds in the programming period 2014-2020. In turn, a region assigned with the value 0 (zero) was not subject to the treatment or, in other words, did not receive EU funds. For this reason - and summing-up -, we here refer to this variable as the ´treatment dummy variable`, according to which regions are categorized as treated (value one) or untreated (value zero), depending on whether they received the treatment (EU funds) or not.

The evidence for the categorization of regions as treated or untreated – and thus their treatment condition - is derived from analyzing European Commission official documents, as well as regulations regarding the 2014-2020 programming period.

Despite the above referred, absolute precision in our analysis demands for an important note at this stage. Hence, as described in section 2.2, within the scope of the cohesion policy, ESIF are available and distributed through all regions across the EU under with different goals. As we were able to explain, over the years, the majority of the financial resources has been allocated to the first objective of the cohesion policy, which in the programming period 2014-2020 was framed as ´Investment in Growth and Jobs`, and within which the ´less developed regions` (with a GDP per capita PPS below 75% of the EU average) were the main beneficiaries. Nevertheless, with considerably lower intensity, the wealthiest regions did also receive a slice of the available funds, including under the ´Investment in Growth and Jobs` goal. This was the case for the ´more developed regions` (with a GDP per capita PPS above 90% of the EU average), but also for the so-called ´transition regions` (with a GDP per capita PPS between 75% and 90% of the EU average, which in the past had lower levels of development, but that are nowadays catching up with the wealthiest ones, and which received transitional support with that purpose).

As a consequence, our comparison spots ´less developed regions` - the ´treated` units, recipients of a considerable bigger slice of ESIF – against ´more developed` and ´transition` regions – the ´untreated` units, recipients of a substantially lower amount of funds. As Pellegrini et al. (2013) assertively pointed out, what is really at stake is a comparison between “hard financed” regions to “soft-financed” ones. For our empirical analysis, this means that the relevant control group (composed of ´untreated` regions) did also receive some form of treatment, though in strongly less significant terms. This fact might induce a downward bias in the final estimations to be reached in the present study, as the units composing the control group (the untreated regions) did also slightly suffer from the effects of ESIF.

d) **Assignment variable – initial level of GDP per capita in PPS**

As described in section 4, the assignment variable of the case is the initial level of GDP per capita in PPS of each region. This value is given by the average regional GDP per capita in the period from 2007 to 2009, when the eligibility to be awarded with funds in the programming period 2014-2020 was defined.

As previously explained, the assignment variable is the only reason determining whether a given region was treated by EU funds or not during the programming period 2014-2020. Those regions with a initial level of GDP per capita below 75% of the EU average were classified as less developed and therefore hardly treated with ´Investment for growth and jobs` funds. In opposition, those with initial levels of GDP per capita above that threshold were not. Therefore, the assignment variable plays a crucial role in the context of the RD design employed in this study.

The data for the assignment variable ´initial level of GDP per capita in PPS` is derived from Eurostat. The regional database of Cambridge Econometrics/ARDECO was also used here for the purpose of completion of missing observations.

## 5.3. Control variables

a) **Service and agriculture share**

Service share and agriculture share in total employment are used as indicators of regional economic and labor profiles.

The agriculture share is given as the percentage of people employed in agriculture, forestry and fishing activities in the total employment of a given region. Both the total amount of people employed in the sector, as well as the total employment for a given region are given in thousand persons and include all individuals aged fifteen years old and older.

In turn, the service share stands for the same representation in total employment of workers employed in the service sector. More specifically, the following activities (and respective workers) are included in this economic sector: wholesale and retail trade, transport, accommodation and food service activities; information and communication; financial and insurance activities; real estate activities; professional, scientific and technical activities; administrative and support service activities; public administration, defence, education, human health and social work activities; arts, entertainment and recreation; other service activities; activities of household and extra-territorial organizations and bodies. The number of individuals employed in the sector, as well as the amount of total individuals employed is given in thousand persons and included everyone aged fifteen or over years old.

Data for both variables is collected from Eurostat, with some missing observations for the agriculture share derived from the regional database of Cambridge Econometrics/ARDECO.

As will be developed in section 5.4, the information on the labor market composition offered by these variables is first used to compare regions on both sides of the cut-off point given by the ´75% rule`. This exercise commonly used in classic randomized experiments does also fit particularly well in our RDD, as it allows for a verification of whether the treatment and control groups are indeed similar and comparable (Lee and Lemieux, 2010). This exercise is thus inherent to the empirical model applied in this study, and particularly for the non-parametric approach, in which the windows of observations considered will be decreased and increased. Secondly, these variables are included as part of the robustness checks performed in section 7. Here, they are used to check for covariates jumps at the cut-off point, as well as control variables included in the regression analysis to be performed.

b) Employment share

The regional employment share is an additional economic and labor indicator included as control variable in this study.

The employment share is given by the rate of people aged between fifteen and sixty-four years old (15-64) employed in a given region over the total employment of the same region. The data for this variable is derived from Eurostat.

The employment share serves the same purposes as the service and agriculture share. Hence, it is first included in a comparison between treatment and control groups; secondly, it is used as a robustness check to verify for jumps at the cut-off point of covariates; lastly, it is introduced in the regression analysis to control for potential bias of the estimations.

#### c) Higher studies share

In addition to the labor and economic profile, we include a measure of education levels at the regional level, in particular by considering the population attainment rates of higher studies. This measure is given by the percentage of people in the population aged 25-64 having completed any form of tertiary education (including university studies or having attended other types of higher education institutions). The data for this variable is derived from Eurostat.

This variable offers a proxy of an important face of a regional development status, namely, the education levels of a given region. It also provides an indication of how qualified the labor force of a particular region is. Similarly to the previous variables, the higher studies share is also used for checking for the comparability of treatment and control groups, as well as in the robustness checks to be performed and as a control variable included in the regression analysis.

#### d) Average annual population

Lastly, regional average annual population is included with the same purposes as mentioned before. This variable gives the average annual population of a region given in thousand persons.

The data for this variable is also collected from Eurostat, with the exception made for the missing year of 2020, as well as for data of UK regions in all the years considered in the analysis, in which cases it is derived from the regional database of Cambridge Econometrics/ARDECO.

## 5.4
## Descriptive Statistics

As the EU cohesion policy aims at delivering economic growth and promoting convergence between EU regions, a first interesting look to be devoted is that of considering how development discrepancies are effectively real across the EU. *Table 1* offers a cross-country picture of the variation of average regional GDP per capita in PPS levels for the EU 28 Member-States, in 2013, just before the starting of the programming period 2014-2020. It also presents a within-country comparison between the maximum and minimum levels of regional GDP per capita in PPS.

| Country | Country Average | Country Max. | Country Min. |
|---|---|---|---|
| Austria | 33765 | 42899 | 23777 |
| Belgium | 30101 | 56645 | 19465 |
| Bulgaria | 10614 | 19398 | 7466 |
| Croatia | 15589 | 16054 | 15123 |
| Cyprus | 21913 | 21913 | 21913 |
| Czech Republic | 22343 | 49440 | 16260 |
| Denmark | 30836 | 42321 | 22817 |
| Estonia | 19908 | 19908 | 19908 |
| Finland | 30696 | 37472 | 24563 |
| France | 23172 | 47610 | 7808 |
| Germany | 31322 | 54115 | 21919 |
| Greece | 16508 | 25279 | 13082 |
| Hungary | 15551 | 28083 | 10872 |
| Ireland | 31650 | 40789 | 21430 |
| Italy | 25954 | 41656 | 15880 |
| Latvia | 16436 | 16436 | 16436 |
| Lithuania | 19277 | 19277 | 19277 |
| Luxembourg | 68711 | 68711 | 68711 |
| Malta | 23249 | 23249 | 23249 |
| Poland | 15889 | 28020 | 12420 |
| Portugal | 19681 | 27665 | 16705 |
| Romania | 14702 | 33489 | 8940 |
| Slovakia | 24810 | 50825 | 13763 |
| Slovenia | 21885 | 25838 | 17932 |
| Spain | 22625 | 32213 | 16605 |
| Sweden | 31640 | 46486 | 27165 |
| The Netherlands | 32961 | 44664 | 25241 |
| United Kingdom | 26670 | 83484 | 17268 |

***Table 1.*** *The cross-country and within country disparities of levels of GDP per capita.* The table shows the average GDP per capita in PPS for each country of the EU-28 (2013) in the second column. The third column displays the maximum regional level of GDP per capita in PPS within each country. The fourth column has the minimum regional GDP per capita in PPS within each country.

The cross-country disparities are easily notable. With an average regional GDP per capita in PPS of around 68711 euros, Luxembourg scores the highest level among EU Member-States, as opposed to the 10614 euros registered in Bulgaria, a number more than six times lower to that registered in the wealthiest country. The economic differences are overall observable between the block of northern and central European countries (including the Scandinavian states, Germany or Austria), performing better than the group of southern Member-States (like Greece, Portugal or Spain) and even more than some of the most recent Member-States from the east and Balkans (such as Latvia, Bulgaria, Romania or Croatia).

The regional disparities are also common when looking at within-country levels of income per capita. Here, even in the wealthiest EU economies the differences are considerable. Hence, for instance, in Austria – which performs second in national averages, following Luxembourg (which has only one region and therefore no regional variation) – the least developed region scores a GDP per capita (PPS) of 23777 euros, just slightly above half of the income per capita of the most developed region in the country. In the extreme opposite side, the most developed region in Bulgaria has an income per capita level which is more than double of that of the least developed Bulgarian region. Interestingly, the Bulgarian`s wealthiest region performs worse than the most lagged region in Austria, which again corroborates a picture of regional disparities in the EU.

Across all the EU 28 (2013) Member-States, the region with the highest GDP per capita (PPS) was Inner London, in the United Kingdom, with an average income of around 83484 euros per inhabitant. The lowest value was found on the Bulgarian region of Severozapaden, with 7466 euros per person.

Following this overall picture of how regional disparities take place across the EU, we now move to provide some indications as to how regions treated by EU funds compare with those untreated ones. In *Table 2*, the group of less developed

| | Treated Regions | Untreated Regions |
|---|---|---|
| **Nr. Regions** | 66 | 201 |
| | | |
| **2007-2009** | | |
| GDP per capita (PPS) | 13603 | 27588 |
| %EU average | 56 | 113 |
| | | |
| **2013** | | |
| GDP per capita (PPS) | 14889 | 28693 |
| %EU average | 59 | 114 |
| | | |
| **2020** | | |
| GDP per capita (PPS) | 17978 | 31178 |
| %EU average | 65 | 112 |

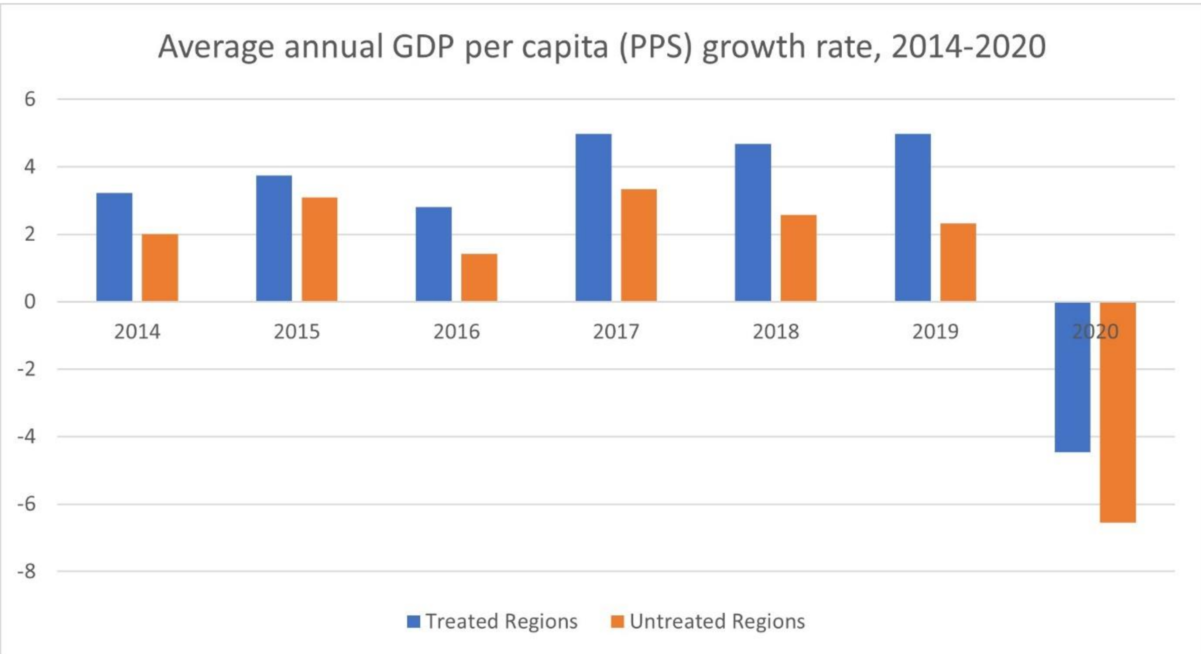
*Table 2. GDP per capita (PPS) of treated and untreated regions in three periods: 2007-2009, 2013 and 2020.* The table shows the average GDP per capita (PPS) of both groups of treated and untreated regions separately considered in 2007-2009, 2013 and 2020. The table also displays the weight of each group´s levels of GDP per capita on the overall average EU GDP per capita in each period considered. The value for GDP per capita (PPS) is given in euros.

regions (treated regions) is compared with that of transition and more developed regions (untreated regions) in their income per capita levels. This comparison is performed for the years eligibility was defined (2007-2009), for one year before the implementation of the programming period 2014-2020 (2013), and for the last year of the programming period (2020). *Table 2* also includes the percentage weight of the average GDP per capita of each group when compared against the average GDP per capita of the entire EU for the respective year.

As the information on the table portrays, the average GDP per capita (PPS) for the 66 treated regions is considerably lower than that of the 201 untreated regions in all years represented. Both groups of regions saw their levels of income per capita rise throughout the years. If we pick the year of 2013 – just before the starting of the implementation of the programming period 2014-2020 – and compare it to 2020 – the last year of the programming period - we notice, however, that the average GDP per capita of the less developed regions increased way more than that of the transition and more developed regions together. Hence, the first group saw a rise of 4375 euros per person, whereas the second faced a 2485 increase. Moreover, when plotted against the average EU GDP per capita in PPS, the mean income per capita of the 66 less developed regions seemed to be converging with their more developed counterparts. Whilst in 2007-2009 – when eligibility for the programming period 2014-2020 was determined – the average GDP per capita of less developed regions represented just 56% of the EU average, by the end of the refered programming period this weight was already around 65%. In contrast, the mean GDP per capita of transition and more developed regions was 113% of the EU average in 2007-2009, a value that dropped to 112% by the end of the implementation of the programming period.

Next, we look at the regional GDP per capita (PPS) growth rates from 2014 to 2020 of both treated and untreated regions. This is a good point to recall that the main goal of interest in this study is to draw conclusions on the impact of ESIF on GDP per capita growth of the less developed regions in the EU.

*Graph 1* portrays a comparison between average annual growth rates of regional GDP per capita (PPS) of treated and untreated regions. The figure shows



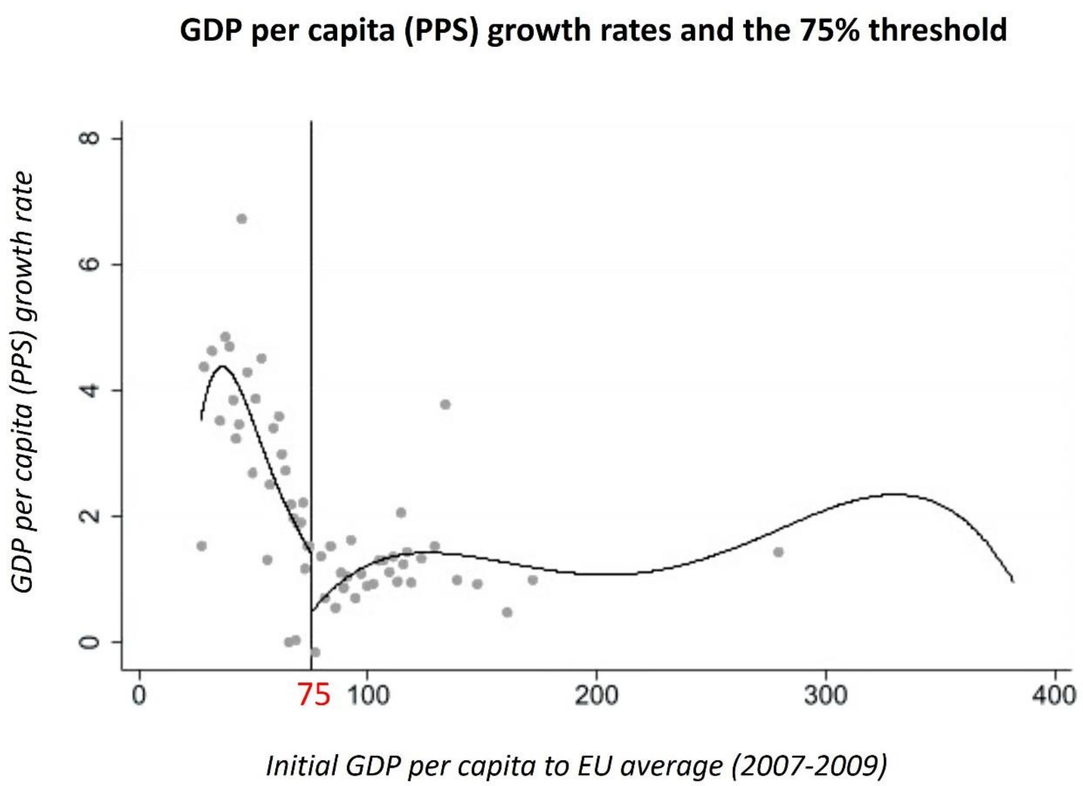

***Graphic 1.*** *Yearly GDP per capita (PPS) growth rate of treated and untreated regions in 2014-2020.* The graphic displays the average yearly GDP per capita (PPS) growth rates of treated and untreated regions considered separately. On the x axis, the year and bars for each group of regions. On the y axis, growth rate in %.

that in all the years considered (from 2014 to 2020), the treated regions revealed to be growing at a faster pace when compared to the group of untreated regions.

Furthermore, the patterns of growth rates of both groups seem to be similar, i.e., they tend to either increase or decrease simultaneously each year. In 2020, all regions felt a drop in their GDP per capita (PPS) growth rates, and all suffered from the economic downturn caused by the COVID-19 outbreak; nevertheless, even here less developed regions seemed to have performed better than the others, as they experienced a lower decrease in their income per capita levels in relation to the previous year. We speculate that this could reveal that regions across the EU suffered similarly from the same economic events during each year, but that the less developed ones – for reasons to be clarified – were still able to grow faster.

The most noteworthy point is, however and again, that less developed regions perform better in all years considered. This might indicate that indeed ESIF can be delivering some form of positive contribution to a raise of income per capita levels, since those regions which are hardly financed by these EU transfers are also the ones having their GDP per capita growing the most. Nevertheless, this remains to be tested in the next section.

A very particular way of visualizing a comparison between treated and control groups in a RD design lies on a graphical representation of the outcome variable plotted against the assignment variable that determines the treatment condition (Lee and Lemieux, 2009; Cattaneo et al., 2019). *Graph 2* performs this exercise, by plotting the annual GDP per capita growth rates of regions against their initial level of GDP per capita compared to the EU average in the years eligibility for treatment was defined. The data is plotted in a different number of bins for each side of the cut-off point, which are constructed using a quantile-spaced portioning scheme for their positioning, and a mimicking variance method to reach the optimal number to be used (Calonico et al., 2014; Calonico et al., 2017). The combination of these methods has the advantage of allowing a better perception of the variability of the data (Cattaneo et al., 2019). Then, flexible polynomial regressions are run




***Graphic 2.*** *The discontinuity at the 75% cut-off.* The graphic plots the discontinuity found at the 75% threshold between treated and untreated regions for the outcome variable GDP per capita (PPS) growth. The graphic represents a 4th order polynomial function. The observations are represented in equally sized bins on both sides of the cut-off.

separately on both sides of the cut-off point for the entire sample. The cut-off point given by the ´75% rule` is traced by the vertical solid line at that value.

This exercise is inherent to the RD design employed in this study (Lee and Lemieux, 2009; Jacob et al., 2012). Hence, in *Graph 2* all points falling to the left of the threshold given by the ´75% rule` stand for GDP per capita growth rates of less developed regions, which by having an initial GDP per capita below 75% of the EU average were assigned with the treatment (the EU funds). In turn, those points to the right of the cut-off point represent growth rates of transition and more developed regions (untreated units), whose condition of having an initial GDP per capita above 75% of the EU average denied them the eligibility for receiving the treatment.

*Graph 2* provides a clear visual perception of how treated and untreated regions compare when plotted against the variable that determined whether they were assigned with the treatment or not. Equally to *Graph 1*, *Graph 2* illustrates that treated regions had overall higher annual growth rates of their GDP per capita (PPS).

More importantly, however, *Graph 2* offers a visual corroboration of the discontinuity of the outcome variable ´annual GDP per capita (PPS) growth` at the cut-off point (Lee and Lemieux, 2009; Jacob et al., 2012), revealed by the jump taking place at the threshold offered by the ´75% rule`. In the absence of the discontinuity observed in *Graph 2*, it would be unlikely that any effect would be reached at the regression analysis (Lee and Lemieux, 2009). Moreover, the size of the discontinuity may also suggest some indications as to the size of the treatment effect itself (Lee and Lemieux, 2009). Nevertheless, this remains to be tested in our regression analysis.

Also relevant in the context of the RD design employed in this study is to perform a comparison between treated and untreated regions for several characteristics that could potentially affect the final results. This comparison is particularly useful for the non-parametric approach that will be used after the parametric estimation. As described in section 4, the assumption taken when performing a non-parametric estimation is that by narrowing down the number of observations solely to those around the cut-off, it is likely that regions with very similar features are compared, and that only the treatment condition differs between them. Bearing this in mind, it is important to test for whether the regions in the narrower windows of observations around the threshold are effectively more comparable (because more equal in all other potential covariates). *Table 3* provides the results of this exercise, performed over four bandwidth sizes: the entire range of observations, regions in the interval 55%-95%, 60%-90% and 65%-85%.

| | All | | 55%-95% | | 60%-90% | | 65%-85% | |
|---|---|---|---|---|---|---|---|---|
| | *Treated Regions* | *Untreated Regions* | *Treated Regions* | *Untreated Regions* | *Treated Regions* | *Untreated Regions* | *Treated Regions* | *Untreated Regions* |
| *Nr. Regions* | 66 | 201 | 37 | 69 | 32 | 52 | 26 | 29 |
| *Average annual Population (Thousand persons)* | 1775 | 1927 | 1851 | 1575 | 1923 | 1523 | 1749 | 1569 |
| *Service Share (%)* | 60 | 74 | 65 | 73 | 65 | 63 | 66 | 72 |
| *Agriculture Share (%)* | 10 | 3 | 8 | 5 | 8 | 5 | 8 | 6 |
| *Employment Share (%)* | 61 | 69 | 61 | 65 | 60 | 65 | 60 | 65 |
| *Higher Education Share (%)* | 23 | 33 | 24 | 31 | 24 | 30 | 24 | 29 |

*Table 3. Regions compared in other variables in different bandwidths.* The table compares treated and untreated regions for several covariates using different bandwidth sizes.

The information displayed in *Table 3* confirms that, indeed, by constraining the regions considered to those closer and closer to the vicinity of the cut-off, they increasingly resemble each other in the variables presented. Hence, for instance, whereas there is a difference of 14 percentage points in the service share when all observations are considered, only 6 percentage points separate treated and untreated regions in the narrowest window of observations (65%-85%).

Lastly, *Table 4* provides summary statistics for the outcome variables, as well as for the covariates used in this study.

| **Variable** | **Observations** | **Mean** | **Std. dev.** | **Min** | **Max** |
|---|---|---|---|---|---|
| GDP per capita (PPS) growth | 1,869 | 1.586393 | 4.586342 | -16.937 | 88.457 |
| Employment growth | 1,869 | .9147822 | 3.602603 | -23.471 | 73.147 |
| Service Share | 1,869 | 70.10428 | 10.06771 | 30.58 | 92.91 |
| Agriculture Share | 1,869 | 4.86824 | 6.095175 | 0 | 46.61 |
| Employment Share | 1,869 | 67.47798 | 8.730908 | 32.2 | 83 |
| Higher Studies Share | 1,862 | 30.95217 | 9.613803 | 11.4 | 68.3 |
| Average Annual Population | 1,869 | 1889.809 | 1582.488 | 28.8 | 12317.56 |

*Table 4. Summary Statistics.* The table provides summary statistics for the outcome variables and other control variables included in the analysis. Information on the higher studies share for the region of Martinique is missing.

# 6.

# Results

The results of the parametric approach are summarized in *tables 5* and *6* (page below). The first table – *table 5* – displays the estimations reached on the basis of the application of a simple pooled OLS. The second table – *table 6* – shows the results achieved with a fixed effects model. The results of the pooled OLS model will be first analyzed, followed by those of the fixed effects.

Within the pooled OLS model, the distinct columns of Table 5 stand for five separate sub-models. Each of these sub-models apply different polynomial orders of the functional form of the assignment variable. The first column for each outcome variable considered displays the results when the linear function of the assignment variable is implemented. Columns 2 to 5, in turn, show the results reached with higher polynomial orders employed as functions of the assignment variable. Polynomials of up to the $5^{th}$ order are included in the analysis. Moreover, for each sub-model performed with a different polynomial order, table 5 shows the respective Akaike Information Criterion (AIC). As stated before, the AIC is used with the purpose of assessing the goodness of fit of the different sub-models applied.

The empirical findings suggests that there is a positive and statistically significant effect of the treatment on the outcome variable ´GDP per capita (PPS) growth` in all the five sub-models applied when using pooled OLS.

That positive effect is bigger for the sub-models applying lower polynomial orders of the functional form of the assignment variable. Hence, when a linear specification is implemented, the treatment is estimated to deliver around 1,6 percentage points higher GDP per capita growth rates for treated regions per year. Moreover, this coefficient is highly significant (at the 1% level). When we turn to a

| | GDP per capita (PPS) growth | | | | |
|---|---|---|---|---|---|
| Polynomial order | *1* | *2* | *3* | *4* | *5* |
| | | | | | |
| **All observations** | | | | | |
| Treatment effect | 1.633*** | 1.392*** | 0.746* | 0.762* | 1.080** |
| (standard error) | (0.311) | (0.387) | (0.442) | (0.441) | (0.484) |
| | | | | | |
| Constant | 1.254*** | 1.940** | 5.267*** | 8.241*** | 11.33*** |
| | (0.402) | (0.769) | (1.343) | (1.757) | (2.613) |
| Nr. Observations | 1869 | 1869 | 1869 | 1869 | 1869 |
| R-squared | 0.025 | 0.025 | 0.030 | 0.034 | 0.035 |
| AIC | 10955.2 | 10954.1 | 10945.0 | 10940.2 | 10937.6 |
| | | | | | |
| | Employment growth | | | | |
| Polynomial order | *1* | *2* | *3* | *4* | *5* |
| | | | | | |
| **All observations** | | | | | |
| Treatment effect | 0.820*** | -0.342 | 0.0117 | 0.0143 | -0.0616 |
| (standard error) | (0.245) | (0.302) | (0.345) | (0.345) | (0.378) |
| | | | | | |
| Constant | -0.915*** | 2.385*** | 0.565 | 1.023 | 0.284 |
| | (0.316) | (0.599) | (1.048) | (1.373) | (2.043) |
| Nr. Observations | 1869 | 1869 | 1869 | 1869 | 1869 |
| R-squared | 0.020 | 0.042 | 0.044 | 0.044 | 0.044 |
| AIC | 10061.8 | 10020.4 | 10015.9 | 10017.6 | 10017.4 |

*Table 5. Results of the parametric approach (pooled OLS).* *, **, *** denote statistical significance at 10%, 5% and 1% level, respectively. The different columns stand for different polynomial orders applied.

| | **GDP per capita (PPS) growth** | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **10** |
| **Polynomial order** | ***1*** | | ***2*** | | ***3*** | | ***4*** | | ***5*** | |
| | Time | Country | Time | Country | Time | Country | Time | Country | Time | Country |
| | | | | | | | | | | |
| Treatment effect (as in table 5) | 1.633*** | | 1.392*** | | 0.746* | | 0.762* | | 1.080** | |
| | (0.311) | | (0.387) | | (0.442) | | (0.441) | | (0.484) | |
| | | | | | | | | | | |
| Treatment effect | 1.633*** | 0.0433 | 1.392*** | 0.0276 | 0.746** | 0.205 | 0.762** | 0.228 | 1.080*** | 0.463 |
| (standard error) | (0.223) | (0.424) | (0.277) | (0.474) | (0.315) | (0.517) | (0.314) | (0.517) | (0.344) | (0.559) |
| | | | | | | | | | | |
| Constant | 1.968*** | 1.236* | 2.654*** | 1.298 | 5.981*** | -0.0398 | 8.956*** | 1.597 | 12.05*** | 4.854 |
| | (0.341) | (0.719) | (0.580) | (1.091) | (0.976) | (1.896) | (1.265) | (2.532) | (1.869) | (3.868) |
| Nr. Observations | 1869 | | 1869 | | 1869 | | 1869 | | 1869 | |
| R-squared | 0.503 | 0.099 | 0.503 | 0.099 | 0.508 | 0.099 | 0.511 | 0.100 | 0.513 | 0.100 |
| | | | | | | | | | | |
| | **Employment growth** | | | | | | | | | |
| **Polynomial order** | ***1*** | | ***2*** | | ***3*** | | ***4*** | | ***5*** | |
| | Time | Country | Time | Country | Time | Country | Time | Country | Time | Country |
| | | | | | | | | | | |
| Treatment effect (as in table 5) | 0.820*** | | -0.342 | | 0.0117 | | 0.0143 | | -0.0616 | |
| | (0.245) | | (0.302) | | (0.345) | | (0.345) | | (0.378) | |
| | | | | | | | | | | |
| Treatment effect | 0.820*** | 0.709** | -0.342 | -0.299 | 0.0117 | -0.0513 | 0.0143 | -0.0344 | -0.0616 | -0.0194 |
| (standard error) | (0.236) | (0.343) | (0.291) | (0.380) | (0.332) | (0.414) | (0.332) | (0.414) | (0.365) | (0.448) |
| | | | | | | | | | | |
| Constant | -0.740** | -1.990*** | 2.560*** | 1.933** | 0.740 | 0.0667 | 1.198 | 1.305 | 0.459 | 1.513 |
| | (0.362) | (0.582) | (0.609) | (0.874) | (1.028) | (1.519) | (1.338) | (2.028) | (1.979) | (3.099) |
| Nr. Observations | 1869 | | 1869 | | 1869 | | 1869 | | 1869 | |
| R-squared | 0.091 | 0.044 | 0.112 | 0.062 | 0.114 | 0.063 | 0.114 | 0.064 | 0.115 | 0.064 |

*Table 6. Results of the parametric approach (fixed effects).* *, **, *** denote statistical significance at 10%, 5% and 1% level, respectively. The columns are organized accordingly to the polynomial order applied and whether a time or country fixed effect is tested.

quadratic function, this effect becomes slightly weaker, but still well above 1 percentage point and statistically relevant at the 1% level.

The point estimations for the same outcome variable when 3rd, 4th and 5th polynomial orders are introduced as functions of the assignment variable remain positive and statistically significant. Compared to the lower order polynomial sub-models, however, they become slightly weaker both in size and statistical significance. Thus, when the 3rd and 4th order polynomials are applied, the treatment is estimated to deliver an increase of GDP per capita growth rates of less developed regions of above 0,7 percentage points per year. In the last sub-model (with a 5th order polynomial) this effect goes up to around 1 percentage point.

An analysis of the goodness of fit of the five distinct sub-models suggests that the models of higher polynomial order tend to better capture the behavior of the data. Using the Akaike Information Criterion, the sub-model that better fits the data is the fifth one, as it has the lowest AIC. Under this sub-model and as referred, the treatment effect is estimated to be of around 1 percentage point per year, and the coefficient is statistically significant at the 5% level.

Having said this, when applying the parametrical approach and using pooled OLS, the ´Investment for growth and jobs` funds are found to positively impact GDP per capita growth rates of less developed regions across the EU. Using the best sub-model specification, the funds are estimated to deliver a 1 percentage point higher GDP per capita growth rates of those regions.

When turning to the alternative outcome variable ´Employment growth`, no conclusive evidence is found of any kind of significant treatment effect. As exhibited in table 5, the coefficients reached for sub-models 2 to 5 are all statistically insignificant. Only when the linear specification is applied (sub-model 1) a positive and statistically significant effect is found, with the treatment here delivering a 0,8 percentage increase of employment levels of treated regions. Similarly to the sub-models run for ´GDP per capita (PPS) growth`, however, the linear specification is also the one providing less ground to be trusted, as it has the highest AIC value among the different specifications.

Furthermore, it is noticed that there is no common pattern as to the type of effect delivered by the treatment on employment growth, as whereas for sub-models 1, 3 and 4 the coefficient is positive, its sign turns to negative in the sub-models 2 and 5.

Overall, the results reached with the parametric approach and by running pooled OLS exhibit the following pattern: first, there is consistent evidence of a positive effect of EU funds in GDP per capita (PPS) growth rates of treated regions; second, the findings for employment growth suggest that no significant treatment effect took place on this outcome.

Next, *table 6* summarizes the results of the fixed effects model. Two types of fixed effects were considered in the regression analysis: entity (country) and time (year) fixed effects. In table 6, ten different sub-models are implemented. Each of the sub-models stands for a specific polynomial order used and, within each polynomial order applied, either a country or fixed effects model was implemented. The results for each of the ten sub-models are presented in the several columns of table 6. The fixed effects model is implemented by means of least square dummy variables.

Regarding the main outcome variable ´GDP per capita (PPS) growth`, the coefficients reached when including time fixed effects are in all equal to those found with simple pooled OLS. Indeed, the results of sub-models 1, 3, 5, 7 and 9 remain all positive, statistically significant and of the exact same size as those presented in table 5. The only difference to be outlined is that for models 5 and 7, the coefficient increased its statistical significance, from the 5% to the 1% level in the two cases.

The sub-models incorporating year fixed effects allow for a control of potential heterogeneous time effects between regions. When compared to the original findings of the pooled OLS, the unchanged coefficients show that such heterogeneity did not occur and that, instead, regions suffered equally from the same economic patterns over the years. This conclusion is in line with what had been already observed in the illustration offered by *Graphic 1* (section 5.4).

In straight contrast, the inclusion of country fixed effects in the analysis leads to a drastic change of the size and significance of the coefficients. In opposition to the original findings, the results of all sub-models 2, 4, 6, 8, and 10, become insignificant for the outcome variable ´GDP per capita (PPS) growth`. At a first look, this would suggest that the treatment is not the main driver of changes in growth rates of GDP per capita, but that instead these are the outcome of country-specific circumstances and economic characteristics. Nevertheless, this conclusion does not entirely hold. Effectively, the insignificance of the results is also the mirror of a lack of sufficient variation on the data when controlling for country-fixed effects. A look at Map 1 presented in section 2.2 is elucidative of this. Moreover, an analysis of the estimations for each country corroborates this, as the estimations remain positive and statistically significant for those countries with treated regions.

The application of the fixed effects model for the alternative outcome variable ´Employment growth` yields similar conclusions. In models 1, 3, 5, 7 and 9, the coefficients reached after controlling for year fixed effects remain equal in the sign, significance and size when compared to those reached with pooled OLS. In turn, in models 4, 6, 8 and 10, controlling for country fixed effects still leads to insignificant results. Only in model 2 the result remains significant, but it becomes weaker and now only significant at the 5% level.

Moving forward, *table 7* introduces the results of the extension of the parametric analysis to the estimation of interaction effects between the treatment and the location of regions within specific groups countries. As developed in section 4, the idea here is to assess whether ´Investment for growth and jobs` funds delivered different outcomes when comparing regions located in groups of countries of different geographical position and largely with cultural and economic dissimilarities between them, as well as with overall different accession dates to the EU.

| | GDP per capita (PPS) growth |
|---|---|
| **Treatment x Central and Northern Europe** | -0.530 |
| | (0.728) |
| **Treatment x Eastern Europe and Balkans** | 1.260 |
| | (1.020) |
| **Treatment x Southern/Mediterranean Europe** | 0.465 |
| | (0.895) |
| | |
| Constant | 1.161 |
| | (0.415) |
| Nr. Observations | 1,869 |
| R-squared | 0.049 |

*Table 7. Results of the parametric approach with interaction effects.* *, **, *** denote statistical significance at 10%, 5% and 1% level, respectively.

The findings summarized in *table 7* suggest that there is no difference in the effect of the funds between the groups of countries considered, as all the coefficients turn out to be statistically insignificant. Nevertheless, it is of an interesting note to refer that the point estimates for both Eastern Europe and Balkans, and Southern/Mediterranean Europe are greater than that for the Central and Northern Europe, and that among the first two groups the point estimate for Eastern Europe and Balkans is higher than that for the group of Southern and Mediterranean countries.

Next, we move to the results of the non-parametric approach, which are summarized in *Table 8*.

| | Rectangular | Triangular | Epanechnikov |
|---|---|---|---|
| | | | |
| **65%-85%** | 2.366* | 2.307 | 2.360 |
| | (1.379) | (1.458) | (1.439) |
| **Optimal Bandwidth (≅60%-90%)** | 0.802 | 2.487** | 2.524** |
| | (0.885) | (1.191) | (1.176) |
| **55%-95%** | 1.701* | 2.118** | 2.021** |
| | (1.011) | (1.044) | (1.030) |

*Table 8. Results of the non-parametric approach.* *, **, *** denote statistical significance at 10%, 5% and 1% level, respectively. The columns indicate the type of kernel function applied, between rectangular, triangular and epanechnikov. The rows indicate the bandwidth applied in each estimation. Linear regressions were performed separately on both sides of the cut-off, and the difference between coefficients estimated.

The table displays the estimations reached for three different bandwidth sizes, including the optimal data-driven bandwidth and two ad hoc fluctuations on the number of observations around the cut-off considered. For every bandwidth implemented, the table shows the results for when rectangular, triangular, and epanechnikov kernels are applied.

Overall, the findings of the non-parametric approach sustain the idea that there is a positive and statistically significant effect of the ´Investment for growth and jobs` funds on GDP per capita growth of treated regions. The estimations reached when the optimal bandwidth is applied are particularly suggestive of this: when both the triangular and epanechnikov kernels are used, EU funds are estimated to deliver a 2,5 percentage points higher GDP per capita growth rate for less developed regions each year. When the rectangular kernel is applied, however, no significant effect is found.

Interestingly, all coefficients are statistically significant when the widest bandwidth is applied. These coefficients are also all positive. Hence, the choice of the widest bandwidth yields to an estimation of a positive treatment effect, which varies from 1,7 percentage points when the rectangular kernel is used, to 2,1 percentage points with the triangular kernel. As Pellegrini et al. (2013) suggest, this might indicate that the wider the bandwidth, the bigger the discontinuity. The same conclusion derivates from the opposite scenario: when the narrowest bandwidth is implemented, a significant effect is only found when the rectangular kernel is used.

When compared to the results of the parametric approach, the coefficients of the non-parametric model are overall more expressive in their size. The main conclusion to be drawn is that the non-parametric approach confirms the existence of a positive and statistically significant effect of the treatment on the treated units.

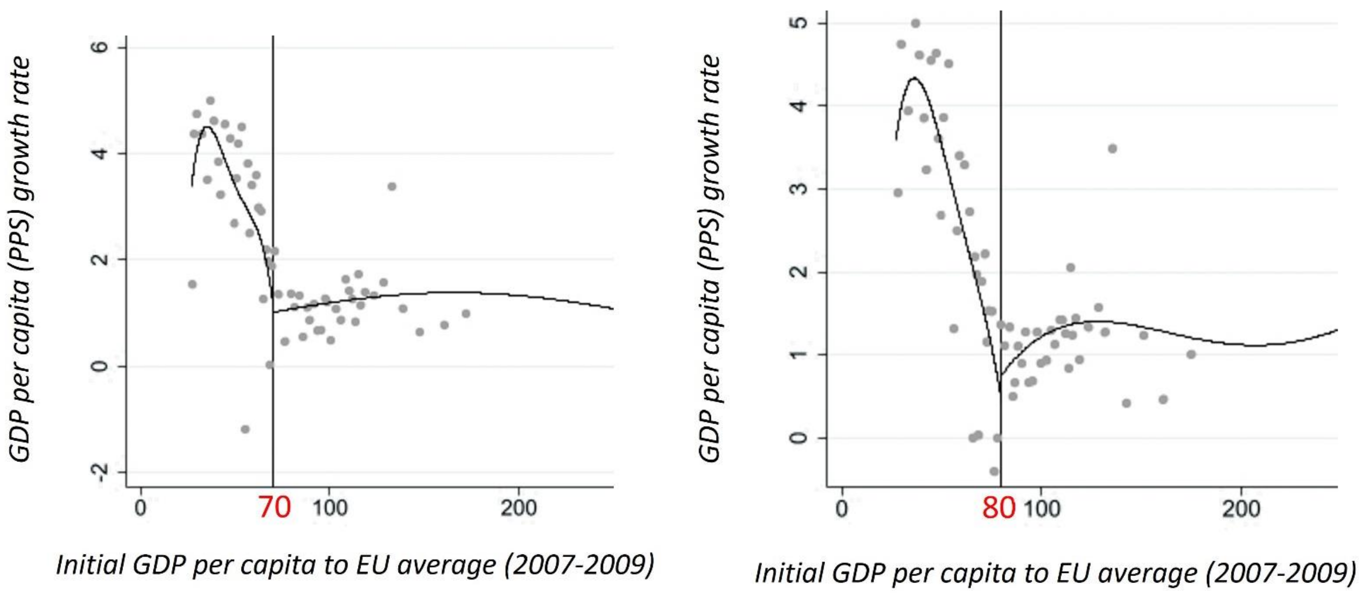
# 7.
# Robustness Checks

In order to check for the reliability of the results outlined in the previous section, several robustness and sensitivity checks are performed. These tests are performed under two different angles: on the one hand, they aim at checking for the credibility of the RD design itself; on the other hand, they are included as specific checks of the validity of the empirical results (Becker et al., 2010). Most of the robustness checks performed in this section are in line with the most influent literature on RD designs, as well as with specific studies on the field (Imbens and Lemieux, 2008; Lee and Lemieux, 2009; Becker et al., 2010).

## 7.1.
## Checks for other jumps at different values of the assignment variable

A first check which is intrinsically connected to the RD design is that of performing a placebo test for the cut-off (Cattaneo et al., 2019). This test aims at confirming that there are no other jumps observed for the outcome variable at different values of the assignment variable rather than that defined as the threshold value (Imbens and Lemieux, 2008; Cattaneo et al., 2019). The presence of different jumps of the outcome variable at points where a discontinuity was not originally expected would imply that the RD setup is wrongly constructed (Cattaneo et al., 2019). In this scenario, in addition to the discontinuity found at the original cut-off point, the jumps at different values of the assignment variable would indicate the existence of additional discontinuities in the relationship between that variable and the outcome. In these circumstances, one would have to consider that potentially there are other treatment effects affecting the outcome, and therefore conclusions could not be taken as to the impact of the original treatment one was interested in (Cattaneo et al., 2019). In simpler words, the results could be biased.

*Graphics 3 and 4 show* the implementation of placebo tests for the cut-off in the present study. Two placebo thresholds are traced at two other values of the assignment variable ´initial level of GDP per capita in PPS`. In both, the treatment condition is assumed to remain unchanged (Cattaneo et al., 2019). The placebo thresholds are thus defined at 70% and 80% of the EU average. *Graphics 3 and 4* provide no indication of relevant jumps of the outcome variable ´GDP per capita growth in PPS` at the two placebo cut-off points.




***Graphics 3 and 4.*** *The placebo tests for discontinuities at other values of the assignment variable.* The graphics plot GDP per capita (PPS) growth against the initial level of GDP per capita when compared to the EU average in the years eligibility for funds in 2014-2020 was defined (2007-2009). The thresholds are represented at different values of the assignment variable, to check for potential jumps. The graphic represents a 5$^{th}$ order polynomial function. The observations are represented in equally sized bins on both sides of the cut-off.

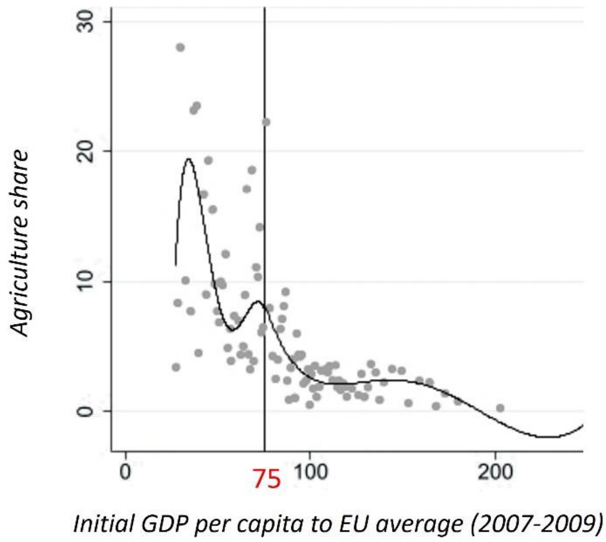
## 7.2.
## Checks for jumps of covariates

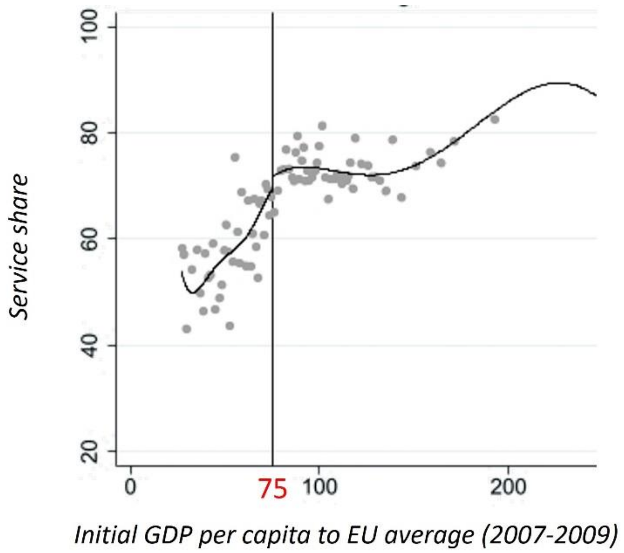
A second standard test of RD designs lies on detecting possible jumps of covariates at the cut-off point (Lee and Lemieux, 2009). This test aims at verifying whether other covariates that are hypothetically related to the outcome variable might also be affecting this last. When a discontinuity is found at the cut-off also for different covariates, one has to consider that it is likely that other factors are also triggering the outcome – and not only the treatment condition under study, as it was originally expected (Imbens and Lemieux, 2008; Pellegrini et al., 2013).

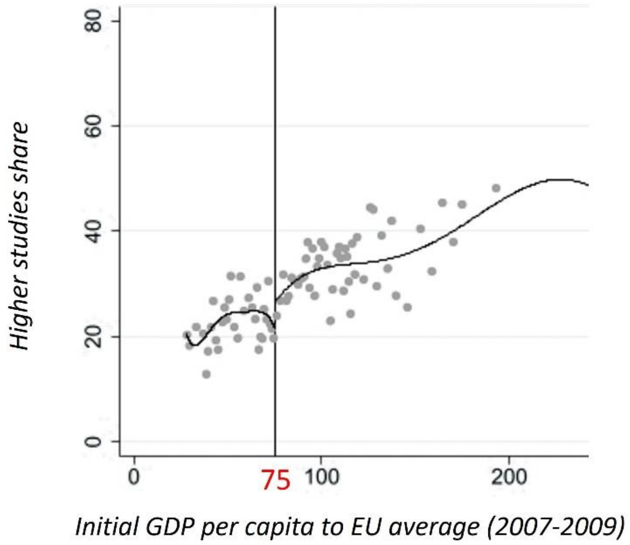
Usually, RD practitioners apply graphical representations of the covariates plotted against the assignment variable, and subsequently try to identify whether there are significant jumps (discontinuities) at the cut-off also for those covariates (Jacob et al., 2012). Using this strategy, Graphics 5 to 9 plot several outcomes of a range of selected covariates (described in section 5) against the assignment variable ´initial level of GDP per capita in PPS`.

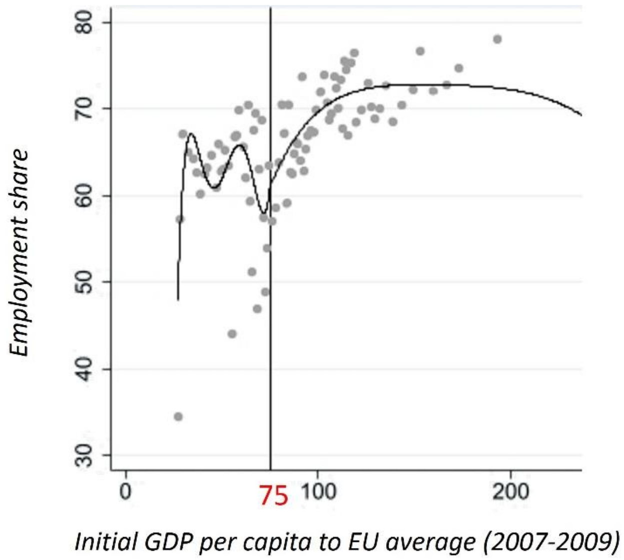
The approach implemented by Becker et al. (2010) and Becker et al. (2018) is followed in what regards the choice of most covariates considered. Hence, we test for potential discontinuities found for variables that mirror the economic profiles of regions, namely: service share, agriculture share, employment share, higher studies share and average annual population. The rationale behind this is that growth rates of GDP per capita might change as the result of distinct, priorly existent regional economic characteristics, instead of due to the implementation of EU funds. If this was the case, the positive results found for the effect delivered by the funds could be biased, and the real explanation of different growth rates between treated and untreated regions could lie on other economic aspects those groups are differently characterized for. By checking for possible discontinuities found for the mentioned variables, however, it is made guaranteed that the real cause of the jump observed at the cut-off is the treatment delivered by EU funds.

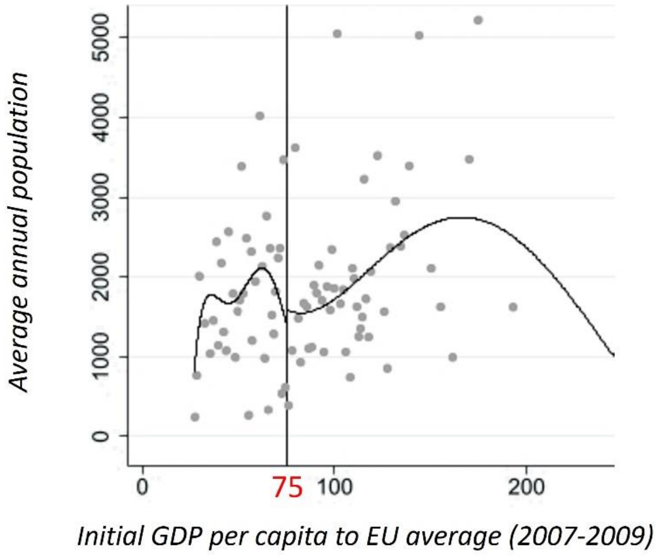
Graphics 5 to 9 show no evidence of significant jumps found for any of the covariates at the cut-off point. In particular, the continuity is perfectly smooth and observable when testing for jumps of the agriculture share, employment share and the average annual population.




***Graphics 5 to 9.*** *Checks for discontinuities in covariates.* The graphics plot several covariates against the initial level of GDP per capita when compared to the EU average in the years eligibility for funds in 2014-2020 was defined (2007-2009). The threshold is at the 75% value. The graphics represent a 5$^{th}$ order polynomial function. The observations are represented in equally sized bins on both sides of the cut-off.

## 7.3.
## Control variables

In addition to RD-specific sensitivity checks, additional robustness checks are performed for the results achieved. As a first of these tests, controls for potential covariates affecting the results are included in the regression analysis. These control variables are those outlined in section 5, which were also previously used to test for possible jumps of covariates: service share, agriculture share, employment share, higher studies share, and average annual population. The justification for the choice of these controls is that they provide an indication of regional economic and labor profiles. If the results achieved in section 6 are unbiased, the inclusion of these controls in the regression analysis should not hamper the estimations. In opposition, a significant change of the estimations necessarily means that other originally unconsidered factors do have an impact on GDP per capita growth rates, and therefore the results of section 6 are biased.

The results achieved after the inclusion of control variables are supportive of the original findings. As observable in table 9, controlling for the indicated variables does not significantly affect the coefficients of the treatment effect in both outcome variables considered.

| | GDP per capita (PPS) growth | | | | | Employment growth | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Polynomial order** | *1* | *2* | *3* | *4* | *5* | *1* | *2* | *3* | *4* | *5* |
| | | | | | | | | | | |
| (All observations) | | | | | | | | | | |
| Treatment effect (as in table) | 1.633*** | 1.392*** | 0.746* | 0.762* | 1.080** | 0.820*** | -0.342 | 0.0117 | 0.0143 | -0.0616 |
| | (0.311) | (0.387) | (0.442) | (0.441) | (0.484) | (0.245) | (0.302) | (0.345) | (0.345) | (0.378) |
| **Control Variables** | | | | | | | | | | |
| Treatment effect | 1.427*** | 1.156*** | 0.550 | 0.550 | 0.899* | 1.050*** | 0.0180 | 0.330 | 0.330 | 0.269 |
| | (0.356) | (0.411) | (0.461) | (0.461) | (0.503) | (0.283) | (0.323) | (0.362) | (0.362) | (0.396) |
| | | | | | | | | | | |
| Service Share | -0.0171 | -0.0158 | -0.0123 | -0.00828 | -0.00421 | 0.0418*** | 0.0466*** | 0.0448*** | 0.0464*** | 0.0457*** |
| | (0.0192) | (0.0192) | (0.0192) | (0.0193) | (0.0194) | (0.0152) | (0.0151) | (0.0151) | (0.0151) | (0.0153) |
| Agriculture Share | 0.0224 | 0.0174 | 0.00391 | -0.00185 | -0.000615 | 0.0239 | 0.00488 | 0.0118 | 0.00959 | 0.00937 |
| | (0.0254) | (0.0257) | (0.0261) | (0.0262) | (0.0262) | (0.0202) | (0.0202) | (0.0205) | (0.0206) | (0.0206) |
| Employment Share | 0.0365** | 0.0389** | 0.0371** | 0.0306* | 0.0297* | -0.000577 | 0.00874 | 0.00970 | 0.00722 | 0.00738 |
| | (0.0171) | (0.0172) | (0.0172) | (0.0174) | (0.0174) | (0.0136) | (0.0135) | (0.0135) | (0.0137) | (0.0137) |
| Higher Studies Share | -0.0231 | -0.0249 | -0.0264* | -0.0272* | -0.0271* | -0.0147 | -0.0216* | -0.0208* | -0.0212* | -0.0212* |
| | (0.0158) | (0.0159) | (0.0159) | (0.0158) | (0.0158) | (0.0125) | (0.0125) | (0.0125) | (0.0125) | (0.0125) |
| Average Annual Population | 0.0000978 | 0.000104 | 0.000105 | 0.0000811 | 0.0000846 | -0.0000666 | -0.0000411 | -0.0000416 | -0.0000508 | -0.0000514 |
| | (0.0000671) | (0.0000673) | (0.0000671) | (0.0000680) | (0.0000680) | (0.0000532) | (0.0000528) | (0.0000527) | (0.0000535) | (0.0000535) |
| | | | | | | | | | | |
| **Controlling for spatial spillover effects** | 1.658*** | 1.351*** | 0.703 | 0.790* | 1.184** | 1.000*** | -0.236 | 0.0937 | 0.108 | -0.00233 |
| | (0.330) | (0.418) | (0.469) | (0.469) | (0.531) | (0.264) | (0.331) | (0.372) | (0.373) | (0.422) |

***Table 9.*** *Original results (parametric) compared to additional robustness checks.* *, **, *** denote statistical significance at 10%, 5% and 1% level, respectively. The columns are organized accordingly to the polynomial order applied.

Hence, when looking at ´GDP per capita (PPS) growth`, the inclusion of these additional controls only slightly decreased the size of the coefficient of the treatment effect. In the linear and quadratic specifications, the coefficient remains significant at the 1% level and the treatment is now estimated to deliver, respectively, around 1,4 and 1,2 percentage points higher GDP per capita growth rates for treated regions. When polynomials of 3$^{rd}$ and 4$^{th}$ orders are applied, however, the coefficient is no longer significant. In model 5 – in which a 5$^{th}$ order polynomial is used – the coefficient is now around 0,9 and remains statistically significant.

The estimations for the treatment effect on the alternative outcome ´Employment growth` are also not significantly affected. Thus, equally to the original results, only when a linear specification is used there is a positive and statistically significant effect of EU funds on employment growth. This effect slightly increases from 0,8 to 1 percentage point. The remaining coefficients for models 2 to 5 remain insignificant but turn now to be all positive.

Overall, controlling for the mentioned variables does not significantly affect the original findings, but instead corroborate the existence of a positive and statistically significant effect of EU funds in GDP per capita growth of treated regions.

## 7.4.
## Controlling for spatial spillover effects

A robustness check is also included in the analysis accounting for potential spatial spillover effects. This test intends to control for the bias generated by the fact that untreated regions might also suffer from some form of treatment, especially due to geographical proximity or interdependences with the treated regions (Becker et al., 2010; Breidenbach et al., 2019). In the case under study, this question is of underlinable importance.

Hence, one has to consider that whilst the treatment under study (EU funds) was delivered to treated regions, these regions are not isolated, and financial and economic flows – especially in the context of a common unified market – take place, leading also to flows of the very effects of implementation of the funds. In this context, it is very likely that neighboring untreated regions to treated ones did also partly experienced the effect of the funds implemented in the nearby territories. Devoting little or no attention to this would lead to results that are potentially contaminated by a downward bias (Becker et al., 2010).

The spatial exclusion approach suggested and implemented by Becker et al. (2010) is here similarly replicated. In this approach, untreated regions which are closely located to treated ones are excluded from the analysis. In the context of this study, only the untreated regions which are direct neighbors with treated regions were excluded from the analysis. After this exercise, the number of observations drops from 267 to 237 regions considered.

The results of the application of this control are displayed in Table 9. All the point estimations achieved after controlling for spatial spillovers remain statistically significant and positive, with the exception made for when polynomial of 3$^{rd}$ order is applied for the outcome variable ´GDP per capita (PPS) growth`, in which case the coefficient loses its significance for a very little margin. However, the findings suggest that, indeed, there might be a downward-bias effect generated by these spillovers, as the coefficients – except that of model 2 - increase in size. In the best specification (model 5), the coefficient increases from 1 to 1,2, which suggests that the original results did not capture a positive effect delivered by the funds also in some untreated regions. This conclusion is to be further discussed in the next section.

# 8.
# Discussion

The empirical evidence confirms Hypothesis 1 of this study and rejects the null Hypothesis 2. Indeed, the final results suggest that the ESIF applied in less developed regions under the ´Investment for growth and jobs` goal contributed to the economic growth of these territories during the programming period 2014-2020. When using a parametric approach, the funds were estimated to have delivered an increase of between 0,9 and 1,2 percentage points in GDP per capita growth rates of less developed regions. In the non-parametric approach, this effect was assessed to be slightly higher, around 2,5 percentage points. These results are in straight line with the findings of some of the most relevant works on the field, including those of Becker et al. (2010) and Pellegrini et al. (2013), which also employed RD designs.

Hence, the original findings of the parametric approach suggested that ESIF funds contributed for a 1 percentage point higher yearly growth rate of income per capita levels on less developed regions. After controlling for potential covariates, the relevant coefficient was around 0,9. Lastly, a point estimate of around 1,2 was reached once a check for eventual positive spatial spillover effects was introduced.

Regarding this last one, as exhibited in table 9, once untreated neighboring regions are excluded from the analysis, the size of the coefficient increased to a contribution of 1,2 percentage points higher economic growth rates in lagging regions (with the significance level remaining the same). This small increase might indicate that indeed there are some spatial spillovers taking place and that funds implemented in treated regions do also have an impact in regions on their proximities. The increase of the point estimate indicates that this impact in neighboring regions is also positive, since once these regions are excluded from the analysis, the effect of switching on the treatment dummy variable (i.e., the effect of the treatment when this variable takes value 1 – effect on treated regions) becomes bigger. If this conclusion holds, the original findings are downward biased, as some effect of the funds is being manifested not only in treated but also in untreated regions, making the comparison less perfect.

The existence of potential positive spillover effects is in line with the findings of Breidenbach et al. (2019), which advanced that such an effect indeed exists and damages the assessment of the impact of the ESIF. Nevertheless, the overall conclusion of a positive effect of the ESIF is not affected by this observation, as what is noticed is an increase rather than a decrease in the coefficient, which also preserves its significance. It remains to further explore whether these spillover effects were captured in their entirety by the model applied in this study. We followed Becker et al. (2010) in applying a spatial exclusion method to perform this robustness check. However, only neighboring regions were removed from the analysis. Future works on the field might want to pave some way as to explore where do these spillovers end and how to more effectively control for their existence.

The several robustness checks performed did also corroborate the positive effect found in the regression analysis.  Thus, both the tests applied on the validity of the regression discontinuity setup, as well as on the results themselves confirmed this conclusion. In particular, the RD strategy employed is valid, as not only the features of the case perfectly fit its application, as after implementing a cut-off placebo test and a check for jumps of other covariates no additional and unexpected significant discontinuity was found. Moreover, to further confirm the initial results, control variables were included in the regression analysis (yielding the point estimate previously mentioned), and the above discussed check for spatial spillover effects was employed.

Moreover, the inclusion of year fixed effects models when running pooled OLS yield to the conclusion that no heterogenous time effect was felt among the regions analyzed. As described in section 6, the inclusion of country fixed effects in the analysis pointed to some uncertainty on whether country-specific characteristics might be more deterministic of the way economic growth evolves for less developed regions. As stated before, however, no precise conclusion could be extracted from this.

In addition to extending the assessment of the impact of ESIF to the programming period 2014-2020, the most original contribution of our analysis also laid on interacting the treatment delivered by ESIF with distinct groups of regions, accordingly to a set of countries with different geographical position and - to a large extent – the same accession date to the EU.

This exercise aimed at verifying whether the effect of ESIF in GDP per capita (PPS) growth of less developed regions varied between the groups of countries included in the analysis. The results exhibit that the effect of ESIF was the same across regions all over the EU in 2014-2020, as the point estimates are not statistically significant for the interaction of the treatment with any of the groups of countries considered. In this way, evidence supporting Hypothesis 3 was also found (and the null Hypothesis 4 rejected), with no differential impact found for the ESIF.

Even though these results shed some light on the potential different impacts of ESIF between distinct groups of regions and countries (having found none for the period 2014-2020 and the categorization of countries applied), a more complete picture with this regard is to be drawn in the future. Hence, it would be a valuable contribution to assess whether there are any country/region-specific structural characteristics affecting the effectiveness of the ESIF, based on other factors or grouping techniques. This exercise – although not entirely original (Ederveen et al., 2016; Crescenzi and Giua, 2020) – is thus one of particular interest for future research on the field. The extension of the analysis performed in this study to previous programming periods might also be valuable.

Lastly, the results of the non-parametrical approach pointed to a positive and statistically significant effect of the ESIF in income per capita growth of less developed regions. The estimations reached on the basis of a non-parametric approach are more expressive than those of the parametric model. When the optimal data-driven bandwidth was applied, the ESIF are estimated to deliver a 2,5 percentage points increase on GDP per capita (PPS) growth rates for the lagging regions.

This result confirms the findings of the parametric approach firstly employed. The increase in the size of the point estimate might be the consequence of restricting the observations considered exclusively to those on the vicinity of the cut-off point. In line with the RD design, this focus on the observations around the threshold leads to a comparison of regions which tend to be very similar in all other aspects, but which differ on whether they were awarded with EU funds or not. This comparison is more perfect than that performed with the parametric approach, in which case all observations were considered, including those further away from the cut-off. A higher point estimate might implicate that some other variable was downward biasing the results of the parametric approach. When only more similar regions were compared - and therefore the potential implication of the effect of a hidden factor -, the real size of the treatment effect was approximated. This does not go without saying, however, that some limitations are also intrinsic to this approach, including the fact that it loses a significant part of the necessary variation to perform the regression analysis.

In any case, the results displayed by the non-parametric approach corroborate the existence of a positive and significant effect of ESIF.

Lastly, the results suggest that no parallel effect significance is found for the alternative outcome variable employment growth. Effectively, no model applied delivered consistent significant estimations of any kind for this variable.

# 9.
# Conclusion

This study was devoted to assessing the effect of the European Structural and Investment Funds – and through them, of the EU cohesion policy – in economic and income per capita growth of less developed regions. In particular, the impact of the funds delivered under ´Investment for growth and jobs` goal during the programming period 2014-2020 was evaluated.

The ´75% rule` used by the EU to categorize regions in different development stages and to establish eligibility to be awarded with the biggest slice of ESIF fitted well in the application of a Regression Discontinuity Design as the methodological approach applied.

The empirical evidence found suggests that in the programming period 2014-2020, ´Investment for growth and jobs` funds had a positive significant impact on GDP per capita (PPS) growth rates of less developed regions. This effect was estimated to be between 0,9 and 1,2 percentage points when a parametrical approach was implemented, and of around 2,5 percentage points with a non-parametrical approach.

Furthermore, this study advanced some conclusions as to whether the effects of the ESIF vary between different groups of regions, accordingly to their geographical position and, largely, accession date to the EU. For the programming period 2014-2020, no different effect was found. In opposition, ESIF did impact treated regions in an equal way.

The validity of the results and of the RDD employed was tested by means of several robustness checks. The inclusion of these tests confirmed that the RDD setup was properly employed and that the results are robust. Only when controlling for spatial spillover effects evidence was found that, indeed, EU funds also delivered positive effects in neighboring untreated regions. This might have slightly downward biased the empirical findings.

Nevertheless, that observation did not affect substantially the answer to our original research question. Hence – and once more – we conclude that ´Investment for growth and jobs` funds contributed to the economic growth of less developed regions.

# REFERENCES

Aiello, F., & Pupo, V. (2012). Structural funds and the economic divide in Italy. *Journal of Policy Modeling*, 34(3), 403–418.

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

Bachtler, J., & Mendez, C. (2007). Who Governs EU Cohesion Policy? Deconstructing the Reforms of the Structural Funds. *JCMS: Journal of Common Market Studies*, *45*(3), 535–564.

Bachtler, J., Mendez, C. & WIshlade, F. (2013). *EU Cohesion Policy and European Integration: the dynamics of the EU Budget and Regional Policy Reform.* Routledge, Taylors & Francis Group, 2016.

Bailey, D., & De Propris, L. (2002). EU Structural Funds, Regional Capabilities and Enlargement: Towards Multi-Level Governance? *Journal of European Integration*, *24*(4), 303–324.

Bailey, D., & Propris, L. D. (2002a). The 1988 reform of the European Structural Funds: Entitlement or empowerment? *Journal of European Public Policy*, *9*(3), 408–428.

Barone, G., David, F., & de Blasio, G. (2016). Boulevard of broken dreams. The end of EU funding (1997: Abruzzi, Italy). *Regional Science and Urban Economics*, 60, 31– 38.

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(9–10), 578– 590.

Becker, S. O., Egger, P. H., & von Ehrlich, M. (2018). Effects of EU Regional Policy: 1989-2013. *Regional Science and Urban Economics*, 69, 143–152.

Benedetto, G. (2019). The History of the EU Budget. European Parliament, Directorate General for Internal Policies. Online available at: http://www.europarl.europa.eu/supporting-analyses

Boldrin, M., & Canova, F. (2001). Inequality and convergence in Europe's regions: Reconsidering European regional policies. *Economic Policy*, 16(32), 206–253.

Breidenbach, P., Mitze, T., & Schmidt, C. M. (2019). EU Regional Policy and the Neighbour's Curse: Analyzing the Income Convergence Effects of ESIF Funding in the Presence of Spatial Spillovers. *JCMS: Journal of Common Market Studies*, *57*(2), 388–405.

Calonico, S., Cattaneo, M. D., & Titiunik, R. (2014). Robust Data-Driven Inference in the Regression-Discontinuity Design. *The Stata Journal*, *14*(4), 909–946.

Calonico, S., Cattaneo, M. D., Farrell, M. H., & Titiunik, R. (2017). Rdrobust: Software for Regression-discontinuity Designs. *The Stata Journal*, *17*(2), 372–404.

Cattaneo, M.D., Idrobo, N., Titiunik, R. (2019). *A Practical Introduction to Regression Discontinuity Designs: Foundations.* Cambridge University Press.

Crescenzi, R., & Giua, M. (2020). One or many Cohesion Policies of the European Union? On the differential economic impacts of Cohesion Policy across member states. *Regional Studies*, *54*(1), 10–20.

Dall'erba, S. & Le Gallo, J. (2008). Regional convergence and the impact of structural funds over 1989–1999: a spatial econometric analysis. *Papers in Regional Science*, 87, 219– 244.

de la Fuente, A., Vives, X., Dolado, J. J., & Faini, R. (1995). Infrastructure and Education as Instruments of Regional Policy: Evidence from Spain. *Economic Policy*, 10(20), 13–51.

Ederveen, S., de Groot, H. L. F., & Nahuis, R. (2006). Fertile Soil for Structural Funds?A Panel Data Analysis of the Conditional Effectiveness of European Cohesion Policy. *Kyklos*, 59(1), 17–42.

Esposti, R., & Bussoletti, S. (2008). Impact of Objective 1 Funds on Regional Growth Convergence in the European Union: A Panel-data Approach. *Regional Studies*, 42(2), 159–173.

European Commission (n.d.). *Available budget of Cohesion Policy 2021-2027.* https://ec.europa.eu/regional_policy/en/funding/available-budget/

European Commission (n.d.a). *History of the policy.* https://ec.europa.eu/regional_policy/en/policy/what/history/

European Commission (n.d.b). *CohesionPolicy 2021-2027.* https://ec.europa.eu/regional_policy/en/2021_2027/

European Commission (2007). Cohesion policy 2007–13: Commentaries and official texts. Luxembourg: Office for Official Publications of the European Communities 2007. Online available at: https://ec.europa.eu/regional_policy/en/information/publications/legislation/2007/cohesion-policy-2007-13-commentaries-and-official-texts

European Commission (2011). Cohesion Policy 2014-2020: Investing in growth and jobs. Luxembourg: Office for Official Publications of the European Communities 2011. Online available at: https://www.europeansources.info/record/cohesion-policy-2014-2020-investing-in-growth-and-jobs/

European Commission (2015). European Structural and Investment Funds 2014-2020: Official texts and Commentaries. Luxembourg: Publications Office of the European Union, 2015. Online available at: https://ec.europa.eu/regional_policy/en/information/publications/legislation/2015/european-structural-and-investment-funds-2014-2020-official-texts-and-commentaries

Federal Ministry of Labour and Social Affairs (2018). 60 years of the European Social Fund – investing in people. Online available at: https://www.bmas.de/EN/Services/Publications/37849-60-years-esf.html

Ferrara, A. R., McCann, P., Pellegrini, G., Stelder, D., & Terribile, F. (2017). Assessing the impacts of Cohesion Policy on EU regions: A non-parametric analysis on interventions promoting research and innovation and transport accessibility: Assessing the impacts of Cohesion Policy on EU regions. *Papers in Regional Science.*

Hagen, T. & Mohl, P. (2009). Econometric Evaluation of EU Cohesion Policy – A Survey. *Center for European Economic Research*, Discussion Paper No. 09-052.

Hahn, J., Todd, P., & Van der Klaauw, W. (2001). Identification and Estimation of Treatment Effects with a Regression-Discontinuity Design. *Econometrica*, *69*(1), 201–209.

Imbens, G. W., & Lemieux, T. (2008). Regression discontinuity designs: A guide to practice. *Journal of Econometrics*, *142*(2), 615–635.

Imbens, G., & Kalyanaraman, K. (2012). Optimal Bandwidth Choice for the Regression Discontinuity Estimator. *The Review of Economic Studies*, *79*(3), 933–959.

Jacob, R., Zhu, P., Somer, MA., Bloom, H. (2012). A practical guide to regression discontinuity. MDRC. Online available at: https://www.mdrc.org/publication/practical-guide-regression-discontinuity

Lee, D. S. (2008). Randomized experiments from non-random selection in U.S. House elections. *Journal of Econometrics*, 142 (2008) 675–697.

Lee, D. S., & Lemieux, T. (2009). Regression Discontinuity Designs in Economics. *Journal of Economic Literature*, *48*(2), 281–355.

Lequiller, F., & Blades, D. (2014). Understanding National Accounts: Second Edition. OECD. Online available at: https://www.oecd.org/sdd/UNA-2014.pdf

Lolos, S. E. G. (2009). The effect of EU structural funds on regional growth: Assessing the evidence from Greece, 1990–2005. *Economic Change and Restructuring*, 42(3), 211–228.

Martins, M. R. & Mawson, J. (1981). Revision of the European Regional Development Fund Regulation. *Built Environment*.

Laia Maynou, Marc Saez, Andreas Kyriacou & Jordi Bacaria (2016). The Impact of Structural and Cohesion Funds on Eurozone Convergence, 1990–2010. *Regional Studies*, 50:7, 1127-1139

Organisation for economic co-operation and development (2009). National accounts at a glance 2009. OECD. Online available at:

https://www.oecd-ilibrary.org/economics/national-accounts-at-a-glance-2009_9789264067981-en

Pellegrini, G., Terribile, F., Tarola, O., Muccigrosso, T., & Busillo, F. (2013). Measuring the effects of European Regional Policy on economic growth: A regression discontinuity approach. *Papers in Regional Science*, 92(1), 217–233.

Regulation 1303/2013 of the European Parliament and of the Council, 17 December 2013. (European Parliament and Council) Online available at: https://eur-lex.europa.eu/legal-content/en/ALL/?uri=celex%3A32013R1303

RodrÍguez-Pose, A., & Fratesi, U. (2004). Between Development and Social Policies: The Impact of European Structural Funds in Objective 1 Regions. *Regional Studies*, *38*(1), 97–113.

Sala-i-Martin, X. (1996). Regional cohesion: evidence and theories of regional growth and convergence. *European Economic Review*, 1325–1352.

Single European Act, 1986. Online available at: https://eur-lex.europa.eu/legalcontent/EN/TXT/?uri=celex%3A11986U%2FTXT

Stiglitz, J.E., Sen, A., Fitoussi, J-P (2009). Report by the Commission on the Measurement of Economic Performance and Social Progress. European Commission. Online available at: https://ec.europa.eu/eurostat/documents/8131721/8131772/Stiglitz-Sen-Fitoussi-Commission-report.pdf

Sutcliffe, J. B. (1995). Theoretical Aspects of the Development of European Community Regional Policy. *Swiss Political Science Review*, *1*(2–3), 1–22.

The World Bank (n.d.). *Metadata Glossary*. Online available at: https://databank.worldbank.org/metadataglossary/all/series

Thistlethwaite, D. L., & Campbell, D. T. (1960). Regression-discontinuity analysis: An alternative to the ex post facto experiment. *Journal of Educational Psychology*, *51*(6), 309–317.

Treaty establishing the European Economic Community, 1957. Online available at: https://eur-lex.europa.eu/eli/treaty/teec/sign

Treaty on the Functioning of the European Union, 2009. Online available at: https://eurlex.europa.eu/legal-content/EN/TXT/?uri=celex%3A12012E%2FTXT

Van Der Klaauw, W. (2008). Regression–Discontinuity Analysis: A Survey of Recent Developments in Economics. *LABOUR*, *22*(2), 219–245.

# DATA SOURCES

## A – Eurostat

Gross domestic product (GDP) at current market prices by NUTS 2 regions
https://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=nama_10r_2gdp&lang=en

Average annual population to calculate regional GDP data (thousand persons) by NUTS 3 regions

https://ec.europa.eu/eurostat/databrowser/view/nama_10r_3popgdp/default/table?lang=en

Employment by sex, age, economic activity and NUTS 2 regions

https://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=lfst_r_lfe2en2&lang=en

Employment rate of the age group 15-64 by NUTS 2 regions

https://ec.europa.eu/eurostat/databrowser/view/tgs00007/default/table?lang=en

Tertiary educational attainment, age group 25-64 by sex and NUTS 2 regions

http://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=tgs00109&lang=en

## B - Cambridge Econometrics/Annual Regional Database of the European Commission's Directorate General for Regional and Urban Policy (ARDECO)

GDP at current prices

https://urban.jrc.ec.europa.eu/trends?lng=en&is=Default&ts=EU&tl=3&dtype=udpp&i=92&db=513&it=metadata&ctx=udp&d=23&cwt=line-chart

Total Population (Regional Accounts)

https://urban.jrc.ec.europa.eu/trends?lng=en&is=Default&ts=EU&tl=3&dtype=udpp&i=66&db=508&it=metadata&ctx=udp&d=20&cwt=line-chart

Total Employment

https://urban.jrc.ec.europa.eu/trends?lng=en&is=Default&ts=EU&tl=3&dtype=udpp&i=121&db=520&it=metadata&ctx=udp

Employment by NACE Sector

https://urban.jrc.ec.europa.eu/trends?lng=en&is=Default&ts=EU&tl=3&dtype=udpp&i=163&db=525&it=metadata&ctx=udp&d=32&cwt=line-chart

```
***MASTER´S THESIS IN PUBLIC ADMINISTRATION - ECONOMICS AND GOVERNANCE***
*STUDENT NAME: Gonçalo Miguel Veiga da Silva
*STUDENT NR.: s2999277
*SUPERVISOR: Eduard Suari-Andreu
*THESIS TITLE:EU cohesion policy: in the name of economic growth? Assessing the impact of ´Investment for growth and jobs` funds in less developed regions in 2014-2020*
*************************************************************************



*Clear / Change of working directory for RDROBUST installation*

clear all
set more off
sysdir
sysdir set PLUS D:\STATA
sysdir

***************************************************************

*RDROBUST Installation*
*This is a special package for Regression Discontinuity Designs, developed by  Calonico, Cattaneo and Titiunik (2014) and Calonico, Cattaneo, Farrell and Titiunik (2017)*It will be used in graphical representation of discontinuity, the computation of the optimal data-driven bandwidth, robustness checks and to perform the non-parametric approach*

search rdrobust

**************************************************************

*UPLOAD OF DATA FILE AND SAVING DATASET*

import excel "C:\Users\Gonçalo Silva\Documents\Gonçalo\LEIDEN UNIVERSITY\MPA\Master Thesis\DATA\DATA for STATA\NEW\Data for STATA.xlsx"

save "C:\Users\Gonçalo Silva\Documents\Gonçalo\LEIDEN UNIVERSITY\MPA\Master Thesis\DATA\DATA for STATA\NEW\Data for STATA.dta"

use "C:\Users\Gonçalo Silva\Documents\Gonçalo\LEIDEN UNIVERSITY\MPA\Master Thesis\DATA\DATA for STATA\NEW\Data for STATA.dta"

**************************************************************

*REMOVING FEW NON-COMPLIERS WITH 75% RULE*//

drop if Region=="Strední Cechy"
drop if Region=="Jihovýchod"
drop if Region=="Kentriki Makedonia"
drop if Region=="Basilicata"
drop if Region=="Alentejo"
drop if Region=="Cornwall and Isles of Scilly"
drop if Region=="Southern Scotland"

save "C:\Users\Gonçalo Silva\Documents\Gonçalo\LEIDEN UNIVERSITY\MPA\Master Thesis\DATA\DATA for STATA\NEW\DATASETwithoutnoncompliers.dta"

************************************************************
*SUMMARY STATISTICS*//

summarize GDPgrowth EMPgrowth service_share agricultue_share employment_share higherstudies_share averageannualpopulation


***********************************************************
***********************************************************
*GRAPHICAL REPRESENTATION OF DISCONTINUITY*

rdplot GDPgrowth EUaverage, c(75) binselect(qsmv)

```

```
*********************************************************
*********************************************************
*ESTIMATIONS*//

*A - PARAMETRIC APPROACH*

*******
    *A.1. - POOLED OLS MODEL*


            *Creating treatment dummy*
gen treatment = 1 if GDPpc0709<18375
replace treatment = 0 if GDPpc0709>18375

            *Regressions*

preserve
reg GDPgrowth treatment GDPpc0709
eststo Polynomial1
reg GDPgrowth treatment c.GDPpc0709##c.GDPpc0709
eststo Polynomial2
reg GDPgrowth treatment c.GDPpc0709##c.GDPpc0709##c.GDPpc0709
eststo Polynomial3
reg GDPgrowth treatment c.GDPpc0709##c.GDPpc0709##c.GDPpc0709##c.GDPpc0709
eststo Polynomial4
reg GDPgrowth treatment c.GDPpc0709##c.GDPpc0709##c.GDPpc0709##c.GDPpc0709##c.GDPpc0709
eststo Polynomial5

esttab,  r2 ar2 se aic star(* 0.10 ** 0.05 *** 0.01)
restore

preserve
reg EMPgrowth treatment GDPpc0709
eststo Polynomial1
reg EMPgrowth treatment c.GDPpc0709##c.GDPpc0709
eststo Polynomial2
reg EMPgrowth treatment c.GDPpc0709##c.GDPpc0709##c.GDPpc0709
eststo Polynomial3
reg EMPgrowth treatment c.GDPpc0709##c.GDPpc0709##c.GDPpc0709##c.GDPpc0709
eststo Polynomial4
reg EMPgrowth treatment c.GDPpc0709##c.GDPpc0709##c.GDPpc0709##c.GDPpc0709##c.GDPpc0709
eststo Polynomial5

esttab,  r2 ar2 se aic star(* 0.10 ** 0.05 *** 0.01)
restore

*********
    *A.2. - FIXED EFFECTS MODEL - Time (year) and Entity (country)

*Performed by means of least square dummy variables*

clear all
use "C:\Users\Gonçalo Silva\Documents\Gonçalo\LEIDEN UNIVERSITY\MPA\Master Thesis\DATA\DATA for
STATA\NEW\DATASETwithoutnoncompliers.dta"
gen treatment = 1 if GDPpc0709<18375
replace treatment = 0 if GDPpc0709>18375

            *Creating country and year dummies*
egen countryid = group( Country)
egen yearid = group(Year)

            *Regressions*

*GDP country*
preserve
reg GDPgrowth treatment GDPpc0709 i.countryid
eststo Polynomial1
```

```
reg GDPgrowth treatment c.GDPpc0709##c.GDPpc0709 i.countryid
eststo Polynomial2
reg GDPgrowth treatment c.GDPpc0709##c.GDPpc0709##c.GDPpc0709 i.countryid
eststo Polynomial3
reg GDPgrowth treatment c.GDPpc0709##c.GDPpc0709##c.GDPpc0709##c.GDPpc0709 i.countryid
eststo Polynomial4
reg GDPgrowth treatment c.GDPpc0709##c.GDPpc0709##c.GDPpc0709##c.GDPpc0709##c.GDPpc0709 i.countryid
eststo Polynomial5

esttab,  r2 ar2 se aic star(* 0.10 ** 0.05 *** 0.01)
restore

*GDP year*
preserve
reg GDPgrowth treatment GDPpc0709 i.yearid
eststo Polynomial1
reg GDPgrowth treatment c.GDPpc0709##c.GDPpc0709 i.yearid
eststo Polynomial2
reg GDPgrowth treatment c.GDPpc0709##c.GDPpc0709##c.GDPpc0709 i.yearid
eststo Polynomial3
reg GDPgrowth treatment c.GDPpc0709##c.GDPpc0709##c.GDPpc0709##c.GDPpc0709 i.yearid
eststo Polynomial4
reg GDPgrowth treatment c.GDPpc0709##c.GDPpc0709##c.GDPpc0709##c.GDPpc0709##c.GDPpc0709 i.yearid
eststo Polynomial5

esttab,  r2 ar2 se aic star(* 0.10 ** 0.05 *** 0.01)
restore

*EMP country*
preserve
reg EMPgrowth treatment GDPpc0709 i.countryid
eststo Polynomial1
reg EMPgrowth treatment c.GDPpc0709##c.GDPpc0709 i.countryid
eststo Polynomial2
reg EMPgrowth treatment c.GDPpc0709##c.GDPpc0709##c.GDPpc0709 i.countryid
eststo Polynomial3
reg EMPgrowth treatment c.GDPpc0709##c.GDPpc0709##c.GDPpc0709##c.GDPpc0709 i.countryid
eststo Polynomial4
reg EMPgrowth treatment c.GDPpc0709##c.GDPpc0709##c.GDPpc0709##c.GDPpc0709##c.GDPpc0709 i.countryid
eststo Polynomial5

esttab,  r2 ar2 se aic star(* 0.10 ** 0.05 *** 0.01)
restore

*EMP year*
preserve
reg EMPgrowth treatment GDPpc0709 i.yearid
eststo Polynomial1
reg EMPgrowth treatment c.GDPpc0709##c.GDPpc0709 i.yearid
eststo Polynomial2
reg EMPgrowth treatment c.GDPpc0709##c.GDPpc0709##c.GDPpc0709 i.yearid
eststo Polynomial3
reg EMPgrowth treatment c.GDPpc0709##c.GDPpc0709##c.GDPpc0709##c.GDPpc0709 i.yearid
eststo Polynomial4
reg EMPgrowth treatment c.GDPpc0709##c.GDPpc0709##c.GDPpc0709##c.GDPpc0709##c.GDPpc0709 i.yearid
eststo Polynomial5

esttab,  r2 ar2 se aic star(* 0.10 ** 0.05 *** 0.01)
restore

******
    *A.3. - CONTROLS AND ROBUSTNESS CHECKS IN THE PARAMETRICAL MODEL*

        *1 - Adding Control Variables*

preserve
reg GDPgrowth treatment GDPpc0709 service_share agricultue_share employment_share
higherstudies_share averageannualpopulation
```

```
eststo Polynomial1
reg GDPgrowth treatment c.GDPpc0709##c.GDPpc0709 service_share agricultue_share employment_share higherstudies_share averageannualpopulation
eststo Polynomial2
reg GDPgrowth treatment c.GDPpc0709##c.GDPpc0709##c.GDPpc0709 service_share agricultue_share employment_share higherstudies_share averageannualpopulation
eststo Polynomial3
reg GDPgrowth treatment c.GDPpc0709##c.GDPpc0709##c.GDPpc0709##c.GDPpc0709 service_share agricultue_share employment_share higherstudies_share averageannualpopulation
eststo Polynomial4
reg GDPgrowth treatment c.GDPpc0709##c.GDPpc0709##c.GDPpc0709##c.GDPpc0709##c.GDPpc0709 service_share agricultue_share employment_share higherstudies_share averageannualpopulation
eststo Polynomial5

esttab,  r2 ar2 se aic star(* 0.10 ** 0.05 *** 0.01)
restore

preserve
reg EMPgrowth treatment GDPpc0709 service_share agricultue_share employment_share higherstudies_share averageannualpopulation
eststo Polynomial1
reg EMPgrowth treatment c.GDPpc0709##c.GDPpc0709 service_share agricultue_share employment_share higherstudies_share averageannualpopulation
eststo Polynomial2
reg EMPgrowth treatment c.GDPpc0709##c.GDPpc0709##c.GDPpc0709 service_share agricultue_share employment_share higherstudies_share averageannualpopulation
eststo Polynomial3
reg EMPgrowth treatment c.GDPpc0709##c.GDPpc0709##c.GDPpc0709##c.GDPpc0709 service_share agricultue_share employment_share higherstudies_share averageannualpopulation
eststo Polynomial4
reg EMPgrowth treatment c.GDPpc0709##c.GDPpc0709##c.GDPpc0709##c.GDPpc0709##c.GDPpc0709 service_share agricultue_share employment_share higherstudies_share averageannualpopulation
eststo Polynomial5

esttab,  r2 ar2 se star(* 0.10 ** 0.05 *** 0.01)
restore


        *2 - Controlling for Spatial Spillover Effects*

*Controlling for spatial spillover effects by means of a spatial exclusion approach of neighboring regions.*
*The code for the creation of a new dataset with excluded regions is presented at the end of this file.*


clear all
use "C:\Users\Gonçalo Silva\Documents\Gonçalo\LEIDEN UNIVERSITY\MPA\Master Thesis\DATA\DATA for STATA\NEW\SpatialSpillovers\Dataset.dta"
gen treatment = 1 if GDPpc0709<18375
replace treatment = 0 if GDPpc0709>18375

preserve
reg GDPgrowth treatment GDPpc0709
eststo Polynomial1
reg GDPgrowth treatment c.GDPpc0709##c.GDPpc0709
eststo Polynomial2
reg GDPgrowth treatment c.GDPpc0709##c.GDPpc0709##c.GDPpc0709
eststo Polynomial3
reg GDPgrowth treatment c.GDPpc0709##c.GDPpc0709##c.GDPpc0709##c.GDPpc0709
eststo Polynomial4
reg GDPgrowth treatment c.GDPpc0709##c.GDPpc0709##c.GDPpc0709##c.GDPpc0709##c.GDPpc0709
eststo Polynomial5

esttab,  r2 ar2 se aic star(* 0.10 ** 0.05 *** 0.01)
restore

preserve
```

```
reg EMPgrowth treatment GDPpc0709
eststo Polynomial1
reg EMPgrowth treatment c.GDPpc0709##c.GDPpc0709
eststo Polynomial2
reg EMPgrowth treatment c.GDPpc0709##c.GDPpc0709##c.GDPpc0709
eststo Polynomial3
reg EMPgrowth treatment c.GDPpc0709##c.GDPpc0709##c.GDPpc0709##c.GDPpc0709
eststo Polynomial4
reg EMPgrowth treatment c.GDPpc0709##c.GDPpc0709##c.GDPpc0709##c.GDPpc0709##c.GDPpc0709
eststo Polynomial5

esttab,  r2 ar2 se aic star(* 0.10 ** 0.05 *** 0.01)
restore


***********************************************************
*B - ESTIMATION INTERACTION EFFECTS - COUNTRIES GROUPED BY GEOGRAPHICAL POSITION*//

clear all
use "C:\Users\Gonçalo Silva\Documents\Gonçalo\LEIDEN UNIVERSITY\MPA\Master Thesis\DATA\DATA for STATA\NEW\DATASETwithoutnoncompliers.dta"
gen treatment = 1 if GDPpc0709<18375
replace treatment = 0 if GDPpc0709>18375

gen GROUP = 0 if Country=="Belgium" | Country=="Denmark" | Country=="Germany" | Country=="Ireland" | Country=="France" | Country=="Luxembourg" | Country=="The Netherlands" | Country=="Austria" | Country=="Finland" | Country=="Sweden" | Country=="United Kingdom"

replace GROUP = 1 if Country=="Bulgaria" | Country=="Czech Republic" | Country=="Latvia" | Country=="Lithuania" | Country=="Hungary" | Country=="Poland" | Country=="Romania" | Country=="Slovenia" | Country=="Slovakia" | Country=="Estonia" | Country=="Croatia"

replace GROUP = 2 if Country=="Portugal" | Country=="Italy" | Country=="Spain" | Country=="Greece" | Country=="Cyprus" | Country=="Malta"


reg GDPgrowth GDPpc0709 i.treatment##i.GROUP


***********************************************************
*C - NON-PARAMETRIC APPROACH*

clear all
use "C:\Users\Gonçalo Silva\Documents\Gonçalo\LEIDEN UNIVERSITY\MPA\Master Thesis\DATA\DATA for STATA\NEW\DATASETwithoutnoncompliers.dta"

*****
        *1-DATA-DRIVEN OPTIMAL BANDWIDTH ESTIMATION*
                *-> Optimal bandwidth calculation using plug-in rule/MSE-optimal rule proposed by Imbens & Kalyanaraman(2012) (cf. Calonico, Cattaneo, Farrell and Titiunik, 2017) *

rdbwselect GDPgrowth GDPpc0709, c(18375) bwselect(mserd) kernel(uniform)

*****

        *2-Regressions (linear) over different bandwidths. The regressions are automatically performed on both sides of the cut-off and the treatment effect is given by the difference on the estimates, which stands for the size of the discontinuity at the cut-off.*

            *a-Optimal data-driven bandwidth on both sides of the cut-off (60%-90%)*

preserve
eststo: rdrobust GDPgrowth GDPpc0709, c(18375) kernel(uniform) bwselect(mserd) all p(1)
eststo: rdrobust GDPgrowth GDPpc0709, c(18375) kernel(triangular) h(3696.677) all p(1)
eststo: rdrobust GDPgrowth GDPpc0709, c(18375) kernel(epanechnikov) h(3696.677) all p(1)

esttab,  r2 ar2 se aic star(* 0.10 ** 0.05 *** 0.01)
restore
```

```
            *b-65%-85%*

preserve
eststo: rdrobust GDPgrowth GDPpc0709, c(18375) kernel(uniform) h(2450) all p(1)
eststo: rdrobust GDPgrowth GDPpc0709, c(18375) kernel(triangular) h(2450) all p(1)
eststo: rdrobust GDPgrowth GDPpc0709, c(18375) kernel(epanechnikov) h(2450) all p(1)

esttab,  r2 ar2 se aic star(* 0.10 ** 0.05 *** 0.01)
restore


            *c-55%-95%*

preserve
eststo: rdrobust GDPgrowth GDPpc0709, c(18375) kernel(uniform) h(4900) all p(1)
eststo: rdrobust GDPgrowth GDPpc0709, c(18375) kernel(triangular) h(4900) all p(1)
eststo: rdrobust GDPgrowth GDPpc0709, c(18375) kernel(epanechnikov) h(4900) all p(1)

esttab,  r2 ar2 se aic star(* 0.10 ** 0.05 *** 0.01)
restore


***********************************************************
***********************************************************

*ROBUSTNESS CHECKS - GENERAL*// (Some added already on previous points)*

    *1 - Checking for jumps at the cut-off for other values rather than the 75% rule*

rdplot GDPgrowth EUaverage, c(80) binselect(qsmv)
rdplot GDPgrowth EUaverage, c(70) binselect(qsmv)

***********

    *2 - Checking for jumps of possible covariates at the threshold*

rdplot service_share EUaverage, c(75) binselect(qsmv) p(5)
rdplot agricultue_share EUaverage, c(75) binselect(qsmv) p(5)
rdplot higherstudies_share EUaverage, c(75) binselect(qsmv) p(5)
rdplot employment_share EUaverage, c(75) binselect(qsmv) p(5)
rdplot averageannualpopulation EUaverage, c(75) binselect(qsmv) p(5)

***********

    *3 - Estimation with other bandwidth choices*

*-> results included previously in non-parametric approach*


***********

    *4 - Comparison of reigons in different badnwidths (included in descriptive statistics)*

        *a-All regions*
sum service_share agricultue_share employment_share higherstudies_share averageannualpopulation if
 EUaverage<75
sum service_share agricultue_share employment_share higherstudies_share averageannualpopulation if
 EUaverage>75

        *b-55%-95%*
drop if EUaverage<55
drop if EUaverage>95

sum service_share agricultue_share employment_share higherstudies_share averageannualpopulation if
 EUaverage<75
sum service_share agricultue_share employment_share higherstudies_share averageannualpopulation if
```

```
 EUaverage>75

        *c-Optimal bandwidth (60%-90%)*
drop if EUaverage<60
drop if EUaverage>90

sum service_share agricultue_share employment_share higherstudies_share averageannualpopulation if EUaverage<75
sum service_share agricultue_share employment_share higherstudies_share averageannualpopulation if EUaverage>75


        *b-65%-85%*
drop if EUaverage<65
drop if EUaverage>85

sum service_share agricultue_share employment_share higherstudies_share averageannualpopulation if EUaverage<75
sum service_share agricultue_share employment_share higherstudies_share averageannualpopulation if EUaverage>75


***********

    *5 - Control variables*

*-> results included previously in parametric approach*


***********
    *6 - Correcting for potential spatial spillover effects*

        *a - generating new dataset with spatial exclusion of neighboring control regions*

clear all
use "C:\Users\Gonçalo Silva\Documents\Gonçalo\LEIDEN UNIVERSITY\MPA\Master Thesis\DATA\DATA for STATA\NEW\DATASETwithoutnoncompliers.dta"


*Drop neighboring control regions*
drop if Region=="Galicia"
drop if Region=="Castilla y León"
drop if Region=="Castilla-la Mancha"
drop if Region=="Andalucía"
drop if Region=="Área Metropolitana de Lisboa"
drop if Region=="Algarve"
drop if Region=="Devon"
drop if Region=="East Wales"
drop if Region=="Lazio"
drop if Region=="Molise"
drop if Region=="Dytiki Makedonia"
drop if Region=="Attiki"
drop if Region=="Sterea Ellada"
drop if Region=="Peloponnisos"
drop if Region=="Ionia Nisia"
drop if Region=="Bucuresti - Ilfov"
drop if Region==" Közép-Magyarország"
drop if Region=="Mazowieckie"
drop if Region=="Zahodna Slovenija"
drop if Region=="Burgenland (AT)"
drop if Region=="Steiermark"
drop if Region=="Kärnten"
drop if Region=="Praha"
drop if Region=="Niederbayern"
drop if Region=="Oberpfalz"
drop if Region=="Oberfranken"
drop if Region=="Chemnitz"
drop if Region=="Dresden"
```

```
drop if Region=="Mecklenburg-Vorpommern"
drop if Region=="Brandenburg"

save "C:\Users\Gonçalo Silva\Documents\Gonçalo\LEIDEN UNIVERSITY\MPA\Master Thesis\DATA\DATA for STATA\NEW\SpatialSpillovers\Dataset.dta"


        *b -> estimations included in estimations of parametrical approach*
```