Measuring the impact of Structural and Cohesion Funds using regression discontinuity design in EU27 in the period 1994-2011

FINAL TECHNICAL REPORT

WORK PACKAGE 14c – Tasks 2 and 3

Ex post evaluation of Cohesion Policy programmes 2007-2013, focusing on the European Regional Development Fund (ERDF) and the Cohesion Fund (CF)

April 2016

EUROPEAN COMMISSION

Directorate-General for Regional and Urban Policy
Directorate B - Policy
Unit B.2 Evaluation and European Semester

Contact: Daniele Vidoni

E-mail: REGIO-B2-HEAD-OF-UNIT@ec.europa.eu

European Commission
B-1049 Brussels

Measuring the impact of Structural and Cohesion Funds using the Regression Discontinuity Design in EU-27, in the period 1994-2011

FINAL TECHNICAL REPORT

WORK PACKAGE 14c – Tasks 2 and 3

Ex post evaluation of Cohesion Policy programmes 2007-2013, focusing on the European Regional Development Fund (ERDF) and the Cohesion Fund (CF)

CCI: 2014CE16BAT098

Directorate-General for Regional and Urban Policy

2016 EN

LEGAL NOTICE

Luxembourg: Publications Office of the European Union, 2016

ISBN 978-92-79-59937-8
doi: 10.2776/3077

Abstract

This ex post evaluation aims to assess the effects of Cohesion policy (CP) on economic growth in the EU-27 regions that have benefited – to various extents – from financial assistance in the programming periods 1994-99, 2000-06 and 2007-13. The study explores how the average effect of CP influences regional GDP growth. The proposed approach is based on a counterfactual causal analysis using a new extension of the Regression Discontinuity Design, that allows us to estimate the non-linear relationship between the intensity of EU regional transfers and per-capita regional growth. We use a new, reliable and comprehensive dataset, stemming from several sources, fully coherent with the CP. The main result is that the positive and statistically significant impact of CP on regional growth is confirmed by the analysis.

Authors of the study

Guido Pellegrini is Full Professor in Statistical Methods for Regional Policies and Statistics at the Department of Social Sciences and Economics, Sapienza, University of Rome (Italy). Previously he was Economist at the Research Department of Bank of Italy, where he managed the Regional Analysis Unit. He participated to several projects and Committees on European and national policy evaluation, in Italy and abroad. He has published several papers in international journals on policy evaluation, industrial and regional policies, economic growth and convergence, local development, economic statistics.

Acknowledgements

The study was prepared by Guido Pellegrini with the collaboration of Augusto Cerqua. The study owes much to Kai Stryczynski, Daniele Vidoni, Terry Ward for many helpful discussions and comments, to the discussants Sascha O. Becker, Guido De Blasio, Santiago Loranca Garcia and seminar participants in Bruxelles, Rome and Ispra (Varese), where preliminary versions of the study were presented. The construction of the final dataset has been the result of a joint work with Daniele Bondonio, Flavia Terribile, Daniele Vidoni and other DG Regio staff members.

Final technical Report on Tasks 2 and 3

Measuring the impact of Structural and Cohesion Funds using the Regression Discontinuity Design in EU-27, in the period 1994-2011

Table of Contents

Table of Figures

1. EXECUTIVE SUMMARY

The aim of this Report is the ex-post evaluation of Cohesion policy impact on economic growth in the EU-27 regions, which have benefited – to various extents – from financial assistance through three programming periods (from 1994 to 2013). The analysis is divided in 3 different tasks. In Task 1 we analyse how varying per capita intensity of EU structural funds affected regional growth in the EU-15 regions in the 1994-2007 period; in Tasks 2 and 3 (this report) we broad the analysis for EU-27 regions, using the available data (1994-2011).

The analysis is focused on both the average impact of European Union Structural and Cohesion Funds (EUF) and the heterogeneity of the treatment intensity. Differently to Task 1, we present here the analysis of the impact of the crisis on effect of EUF and the evaluation of these effects on the new State Members. We use a new regional dataset, which is fully coherent with Structural Funds Regulations, and the spatial grid defined at level 2 of the 2006 NUTS classification. We propose a new method for estimating the effects of intensity on growth, extending the Regression Discontinuity Design (RDD) framework to the case of continuous treatment.

The results basically confirm the positive effects of EUF transfers on regional growth presented in Task 1. In the case of the fully specified model and normalized by the population, the effect captured by the treatment dummy is +1,9 percentage points more in terms of annual GDP growth, when the intensity is on average; this is due to the strong difference in the average treatment between treated and not treated regions. However, this is the effect for the new member state regions. The effect for EU-15 regions is lower, equal to +1.2 percentage points, closer to results presented in Pellegrini et al. (2013), where the average impact is equal to +0.9 in the parametric approach, and the results in Task 1 (+0.7). The findings of this evaluation show that regional policies have a role for stimulating regional growth and economic development.

However, the results are less statistically significant than in Task 1. The main reason is the limited availability of data. We need additional information (beyond the year 2010) for a more robust empirical analysis of the last programming period (2007-2013), where the heterogeneity across regions is higher, due to the presence of new Member States and the largest economic crisis in Europe since WWII was in action. Our results show a higher impact of EUF for the new member state regions, and a lower impact during the crisis.

Another important result is that the estimated marginal impact on regional growth of further increasing the intensity of the EUF tends to be higher on average for the regions that do not already receive a high intensity of the EUF. In other words, the marginal impact on growth of adding more EUF intensity tends to decrease for the regions with high EUF intensity. A great deal of caution, however, should be exerted in interpreting these results as supportive of the hypothesis that diminishing returns to investment and/or limited absorption capacities may be in place to hamper the full economic-development potential of the high intensities of the EUF transfers.

However, GDP is only one dimension on the target of EU regional policy, which is oriented to remove economic and social disparity across European regions. Therefore effects on GDP cannot exhaust the purpose of the regional policy. Moreover, the decreasing of the marginal returns of transfers can be differentiated by target. For instance, the impact could be always positive in term of employment or social inequality. Therefore this analysis can give some suggestion on interregional allocation of EUF, but cannot be the base of a new allocation of EUF transfers among European regions.

1. AIMS AND SCOPE

1.1. Aim of the project

The aim of the full project is the ex-post evaluation of Cohesion policy impact on economic growth in the EU-27[1] regions, which have benefited – to various extents – from financial assistance through multiple programming periods, from 1994-1999 to 2007-2013. The analysis is divided in 3 different tasks. In Task 1 we analyse how varying per capita intensity of EU structural funds affected regional growth in the EU-15 regions in the 1994-2007 period; in Tasks 2 and 3 (this report) we broad the analysis for EU-27 regions, using the available data (1994-2011).

The object of Task 2 and Task 3 is the measurement of the European Union Structural and Cohesion Funds[2] financial transfers (EUF) impact among EU-27 regions considering the heterogeneity in the intensity of transfers across European regions. The study explores how the average effect of EUF is affected by the per capita intensity of European regional funds and influences regional GDP growth in the EU-27 regions. The larger set of information with respect to Task 1 allows us to analyse the impact of the crisis on the effect of the European regional funds and to evaluate these effects on the new Member States. The proposed methodology is based on a counterfactual causal analysis and RDD (a method with high internal validity), that allows us to estimate the non-linear relationship between the intensity of EU regional transfers and the per-capita economic growth. The proposed approach is the same used in Task 1, and it is based on the methodology for explaining effects heterogeneity presented in Becker, Egger and von Ehrlich (2013), modified for taking into account the intensity of the treatment.

1.2. Introduction and motivations

The aim of the study is to evaluate the impact of Structural and Cohesion Funds (EUF) on regional economic growth in Europe The main difference of this study with the copious literature on this issue is twofold: first, we want to verify if the average impact of EUF on regional growth also depends on the heterogeneity of transfers intensity, measured by the normalised amount of funds distributed in each region; second, the evaluation is based on the Regression Discontinuity Design (RDD), a quasi-experimental method with strong internal validity. Like the analysis in Task 1, is the first time that, to our knowledge, the framework of the RDD is extended to the case of continuous treatment.

The effect of EUF is strongly heterogeneous by country, region and time. There are several factors that could affect the impact of EUF on different dimension on regional growth, and the use of counterfactual methods for policy evaluation often captures only average effects, without explaining differences in the outcome of EU regional policy among regions. The high heterogeneity of regional transfer intensity across regions, also within the same country, suggests that the intensity of allocated funds between regions is a primary source of variability of the impact. The differences in the intensity of transfers reflect the choice to allocate more resources to those regions that are particularly in need, to sustain areas with economic and social distress, measured with specific indicators, and finally to maintain some qualitative judgment by EU and individual Members.

---

1 Croatia is not included as it acceded to the European Union only in 2013.

2 The Structural and Cohesion Funds comprise the ERDF, the ESF and Cohesion Fund.

The relationship between the aid intensity and the impact of EUF is not known. Economists and policy makers ignore whether this relationship is linear, that is, if increasing the EUF transfers would proportionally increase the impact on economic growth, or if it takes some other form. In other words, we do not know if the marginal efficiency of transfers, using economic jargon, is constant or in some parts of this relationship it is increasing or decreasing. As mentioned in Task 1, several arguments can justify the presence of a non linearity in the dose-response function of the EUF transfers. Clearly, the assumption of diminishing returns to investment (and to subsidised investments) implies that a larger number of investment projects carried out would be associated with a lower return to investments (or transfers). In this case, after a determined level of EUF transfers no additional (or even lower) per capita income growth effects would be generated (Becker et al., 2012). However, especially in less developed regions, diminishing returns can be at the end of a process dominated by increasing returns if, for example, a network infrastructure is completed. However, the effect of diminishing returns can be different across European regions, depending on the stage of development, the quality and quantity of social capital, and potential demand.

A different reason is the limited absorbing capacity of EUF transfers, especially in less developed countries and regions, which affects the making the most of the investments taking place in their territory. This would imply that some regions use EU transfers increasingly inefficiently as they receive more transfers. Several authors attribute this effect to a lack of administrative capacity. In a recent paper, Rodríguez-Pose and Garcilazo (2015) note that the European Commission (EC) adopted the view that poor institutions can undermine efforts to achieve greater economic cohesion and hinder the effectiveness of regional development strategies, as stated in the EC's Fifth Cohesion Report

Although the literature on the impact evaluation of the EUF is very wide (for a recent review see Pellegrini et al., 2013), only a few papers evaluate the effects of transfer intensity. In Task1 we reviewed Mohl and Hagen (2010), and Becker et al. (2012), both using the GPS approach. They find that, overall, EU transfers enable faster growth in the recipient regions as intended. Moreover, the impact of the crisis on the dimension of EUF effect is basically not explored. We know only a new paper (Becker et al., 2016) that extends the analysis of EUF from 1989 to 2013.The main finding of the paper is that EUF transfers generate additional growth across founded regions, but the effects are weaker when we consider the most recent programming period.

From a methodological point of view, Mohl and Hagen (2010) and Becker et al. (2012) use the GPS approach, a non-parametric method to estimate treatment effects conditional on observable determinants of treatment intensity. The GPS is one of the methods proposed in the literature to address the problem of a continuous treatment. In this case the policy mechanism can be away from an experimental data framework because of the presence of multiple non random selection processes, related not only to the participation decision but also to the treatment level assignment. In presence of continuous treatment the selection bias problem cannot be tackled using the usual estimation method developed for the binary treatment case. GPS is the main econometric technique for policy evaluation in presence of continuous treatment, as it should be able to correct for selection bias into different levels of treatment intensity by comparing units that are similar in terms of their observable characteristics. As we reviewed in Task 1, the literature proposes few matching estimators for continuous treatment. The main approaches are based mostly on two methods: the generalisation of the propensity score approach in a regression contest (Hirano and Imbens 2004; Imai and Van Dijk, 2004), that is used by Mohl and Hagen (2010) and Becker et al. (2012), and the use of matching method with doses (Behrman, Cheng and Todd, 2004; Cattaneo, 2010). However, in both cases the analysis is limited to the treated group, and the comparison between treated and non-treated units is absent. Moreover, there is not an explicit concern on the selection process related to the treatment level assignment. A different two step matching approach to estimate the causal treatment effects as a function of the doses was proposed by Adorno, Bernini and Pellegrini (2007). The estimator matches treatment and comparison units that are similar in terms of their observable characteristics in both the selection processes (the participation decision and the treatment level assignment).

However, all the estimators based on the matching approach suffer the strong heterogeneity of regions, which is hardly captured by the observed covariates. Moreover, none of these papers have properly exploited the source of local randomness due to the sharp discontinuity in the assignment of different transfer intensity (75% of average GDP criterion). A different solution is proposed in our paper by using the continuous RDD, which allows for the first time a compelling evaluation strategy also in presence of a continuous treatment. The RD design was first introduced by Thistlethwaite and Campbell (1960), and the seminal paper by Hahn et al. (2001) provided the crucial identification results.3 The RDD is typically a preferable choice because it exploits local randomness conditions near the cut-off (k) of the intensity assignment rule and does not require to explicitly measure and observe all the pre-intervention regional heterogeneity. However, in the presence of small samples of regions near the cut-off (k) of the intensity-assignment rule, the GPS model can be a very viable option to ensure the balancing of the relevant control variables and the higher efficiency of the estimates.

The main differences of Task 2 and 3 with Task 1 is that in Task 2 and 3 we enlarge the analysis to include the estimates of the EUF impacts on various regional growth outcomes in the EU-27 regions during the three last programming cycles, using the available data (1994-2011) and providing some empirical evidence on the effects of EUF during the crisis. The main vantage is not only that we have a higher number of regions and therefore more observations (263 instead that 208) in the econometric estimation, but basically that we can explore whether the impact of EUF is heterogeneous across the new member states and during the economic crisis. However we have also several disadvantages: The heterogeneity and the confounding factors respect to all EU-27 countries and all programming period are higher that in the EU-15, 1994-2006 sample; the heterogeneity across regions is higher, due to the presence of new member states with a different economic systems with respect to UE-15 regions; the longer period ends with the largest economic crisis in Europe since WWII; moreover, the data availability reduces the analysis of the last programming period to the time span 2007-2010, where the programs financed by EUF have not completed their effect. From this point of view, the analysis of the last period has to be confirmed when the data availability will be broader.

[3] See Lee and Lemieux (2010) for a comprehensive review of the RD design.

2. DATA

2.1. Novelty of data

This study is based on a new, reliable and comparable dataset, stemming from several sources. We use the new dataset for Task 1, 2 and 3, and here we describe the information already presented in the Report on Task 1. The spatial grid used in our work is defined by 263 EU-27 regions at level 2 of the NUTS classification. We use the NUTS 2006 classification with adjustments to include data from all the three past programming period:

- considering the 2003 and 2006 amendments to the NUTS 1999 classification, regions that, from 1999 to 2003 and/or from 2003 to 2006, experienced any "split" are included as in the NUTS 1999 classification (for Germany, Brandenburg; for Spain, Ceuta y Melilla; for Italy, Trentino-Alto Adige);
- NUTS-2 regions that, from 1999 to 2003 and/or from 2003 to 2006, experienced any "merge" are included as in the NUTS 2006 classification (for Germany, Sachsen-Anhalt);
- NUTS-2 regions that experienced, from 1999 to 2003 and/or from 2003 to 2006 any merge and split together, are considered as in the NUTS 2006 classification (some regions in Portugal and Finland);
- Denmark and Slovenia are presented with just one NUTS-2 (as in the NUTS 1999 and 2003 classifications).

Data on EU Structural and Cohesion Funds payments to Member States, broken down by programming period (1994-1999, 2000-2006, 2007-2013) and region per year, has been provided by the European Commission-DG REGIO. The originality and relevance of this dataset arises from its internal coherence (EU payments by operational programme per year) and extensiveness (it covers all the main funds, including the Cohesion Fund, the European Regional Development Fund (ERDF), the European Social Fund (ESF), the European Agricultural Guidance and Guarantee Fund (EAGGF); the Financial Instrument for Fisheries Guidance (FIFG)). Note that only data on the EUF payments are considered, without national or private co-financing of the projects, in order to evaluate the multiplier of EU regional policy. However, national co-financing tends to be proportionate to EU funding and therefore may not change substantially the relative amount of funding going to different regions.

This information was cross-checked and matched with the ESPON and ISMERI regional databases on Structural and Cohesion Fund spending, in order to define the volumes of payments from the EU to the individual NUTS-2 regions. For the programming period 1994-1999, as the totals per country are very similar to the totals produced by ESPON, we apply the ESPON NUTS-2 spending breakdown to define the volumes of payments from the EU to the individual NUTS-2 regions by Member State. For 2000 onwards, for each country, we adopt the regional weights of the ISMERI expenditure breakdown by NUTS2 regions and apply it to the European Regional Development Fund (ERDF) and Cohesion Fund data; for the European Social Fund (ESF), for each Member State we apply a different set of regional weights based on total population; for the European Agricultural Guidance and Guarantee Fund (EAGGF), and the Financial Instrument for Fisheries Guidance (FIFG), the regional weights are based on employment in agriculture.

For the empirical analysis we link these data with information on various regional characteristics from Eurostat and Cambridge Econometrics' Regional Databases. The data cover the years 1989 through 2011/2013.

As main outcome variables of interest we consider the average annual growth rate of per capita real GDP at NUTS-2 level. In order to use a unique source of information, data on regional GDP in volume are taken from Cambridge Econometrics' Regional Database.

We also consider alternative outcome variables and control variables at the level of NUTS-2 regions. Data on the real growth rate of GVA for the period 1994-2011 are available from Cambridge Econometrics. Information on Employment (6 sectors) and Gross Fixed Capital Formation at NUTS2 regional level for the period 1989-2011 are taken from Cambridge Econometrics (consistent with Eurostat data, but more complete). We also employ information on the percentage of 25-64 year-olds with tertiary education from Eurostat (European Union Labour Force Survey) for the period 2000-2013, as well as data on population (total and the share of population aged 65 and over) and population density.

Cambridge Econometrics and the regional database of Eurostat were the main source of other pre-treatment covariates used in the analysis: employment rate, 15-64; productivity (GVA per hour worked); share of employment in service sector; and share of population over 65.

As in Pellegrini et al. (2013), we exclude from analysis some regions: in the first and second periods, like in Task 1, we exclude four NUTS 2 regions whose level of per capita GDP in the period 1988–1990 (i.e., the reference period for the determination of Ob. 1 eligibility by the European Commission) was above 75 per cent of EU average, but were included in Ob. 1 for 'political reasons': Prov. Hainaut (BE), Corse (FR), Molise (IT), Lisboa (PT). Moreover, a few regions were clear outliers in a certain programming period and were dropped from the sample.4 For example, in the first programming period 2 regions were considered outliers in the analysis (Aragon in Spain and West Macedonia in Greece).

In line with the RDD approach, we selected a restricted sample, which includes the regions closest to the discontinuity. In order to still maintain a sufficient number of degrees of freedom, we have eliminated the lowest quarter for treated regions (in terms of initial level of per capita GDP) and the upper quarter for the non-treated regions. Clearly, the number of regions differs across different programming period, and also in the restricted sample.

The samples in the different programming period are the following:

Programming period	Complete dataset		Restricted dataset	
	Treated Reg.	Non Treated reg.	Treated Reg.	Non Treated Reg.
1994-1999	57	150	39	112
2000-2006	53	155	40	113
2007-2013	82	180	60	130

An important question is the normalisation of the EU regional expenditure. A normalisation is needed: The method used by the Commission in the allocation of resources for each member-state is based on a financial allocation per inhabitant per year, to be applied to the population living in the eligible regions (Barbieri and

[4] The criterion for outliers is to have received funds above the average plus 2.5 times the standard deviation of the respective treatment group in the restricted sample (once excluded the lowest quarter for treated regions - in terms of initial level of per capita GDP- and the upper quarter for the non-treated regions; see below in the paper).

Pellegrini, 1999). From the above, the average population by region seems the "natural" normalisation variable. However, in the literature, the beginning-of period GDP has been used (Mohl and Hagen, 2010; Becker et al., 2012). The reason is that this share is a clear minimum target of the impact of EUF on the economy. From our point of view there is not a prevailing method, therefore we used both indicators in our analysis. Therefore the EUF are also normalized for GDP in 1994 for the 1994-2006 period, for GDP in 2007 for the 2007-2013 period.

2.2. Some descriptive statistics and figures

In Figure 1 we present the distribution of EUF intensity by region, sorted by the 1988-1990 per capita GDP (our forcing variable), for the period 1994-1999 and the period 2000-2006. In Figure 2 we show the distribution of EUF intensity for the third period. Clearly the number of the regions and the intensity differ by period. In Figure 1 the regions we excluded from our sample because of 'political reasons' are in orange. The two groups of treated and non-treated regions are clearly differentiated. We graph the EUF intensity, defined as the total amount of Structural Funds in the period 1994-2006 by region, normalised both by population (the population in 1994) and by GDP (the level of GDP at constant prices in 1994).

Figure 1 shows how the excluded regions in green are clearly outliers in their groups, and do not modify substantially the distribution of the variable in both cases. The line of discontinuity accurately identifies the two groups of treated and non-treated (sharp design). There are very few cases where the contribution of EUF is close to zero. The normalisation affects moderately on the differences between the two groups. As expected, the variability of the intensity is slightly lower for the variable normalised with respect to the GDP, especially for the non-treated. Figure 2 shows a slightly different picture. We graph the EUF intensity, defined as the total amount of Structural Funds in the period 2007-2010 by region, normalised both by population (the population in 2007) and by GDP (the level of GDP at constant prices in 2007). The line of discontinuity identifies the two groups of treated and non-treated regions, but the differences are "less sharp" and there are cases where non treated regions receive a EUF intensity higher than some treated regions. Also in this case the normalisation affects moderately on the differences between the two groups.

Figure 1: Structural Funds intensity by NUTS2 regions period 1994-2000

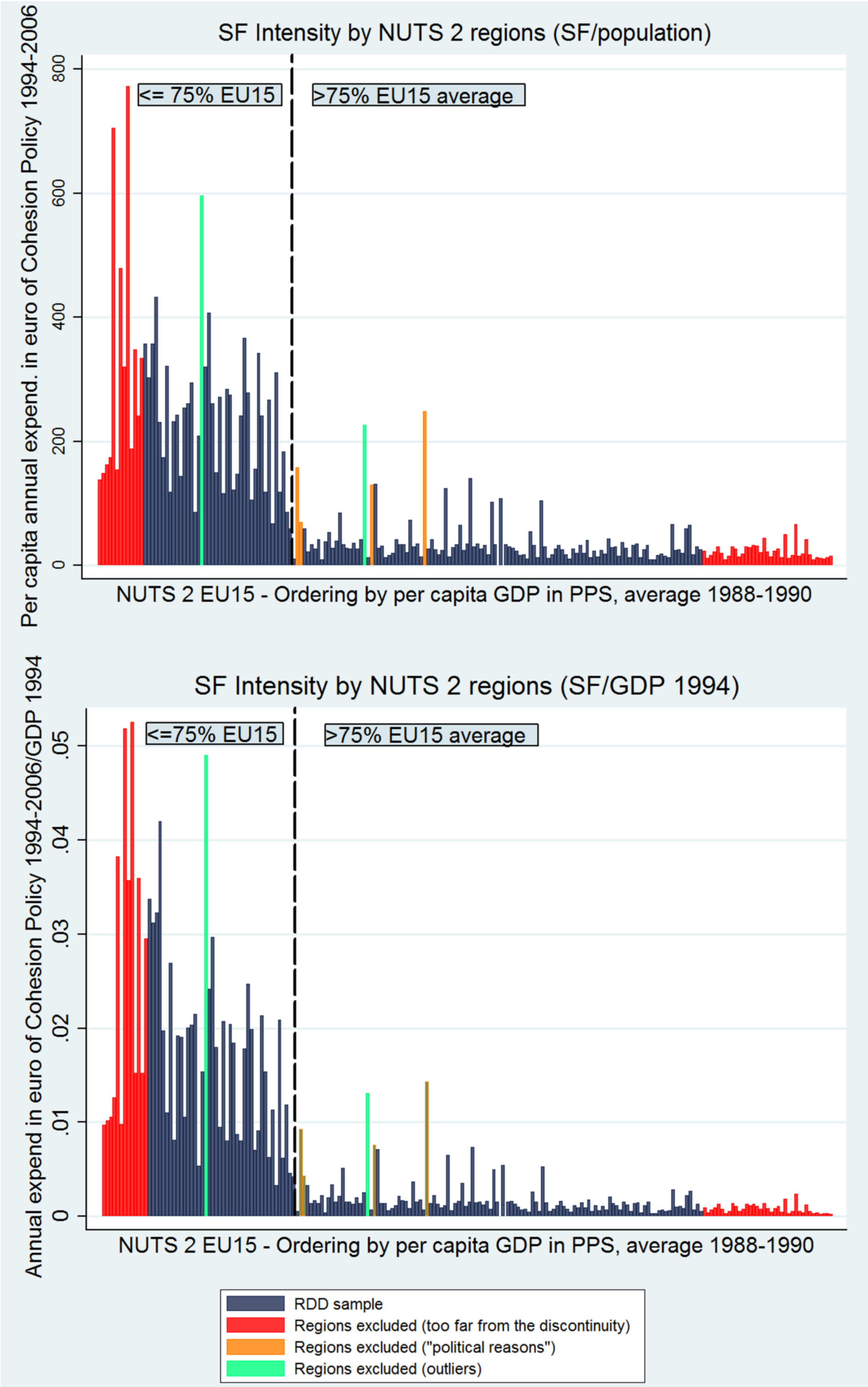
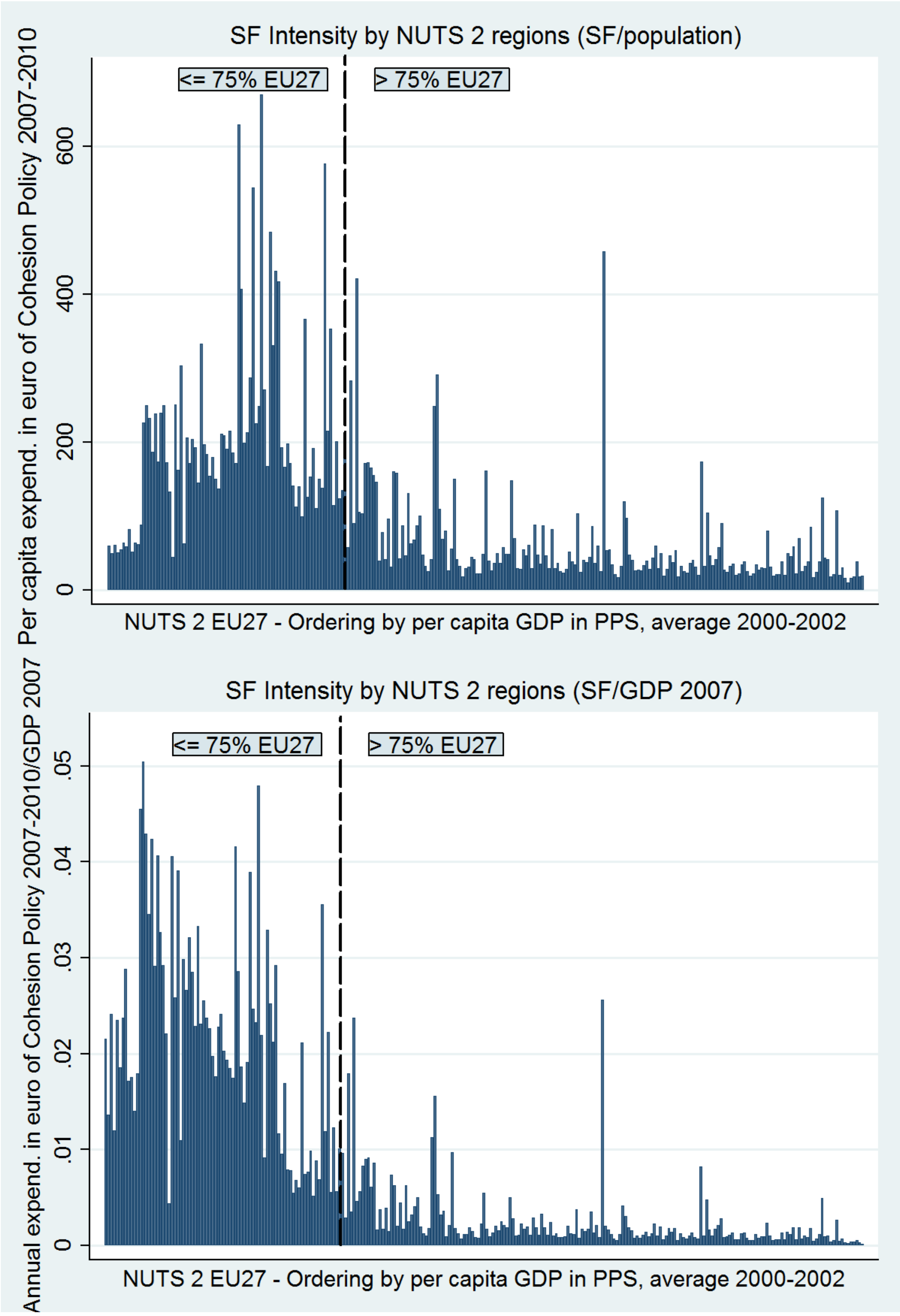


Figure 2 : Structural Funds intensity by NUTS2 regions period 2007-2013



The geographical position of treated and non- treated regions in the EU in the period 1994-2006 and 2007-2013 are presented in Figures 3 and 4: the standard core-periphery picture is clearly exhibited, being the treated region mostly in the periphery. Figure 4 shows that the new member states regions are basically all "treated", and therefore a coherent control sample cannot be easily found.

Figure 3 : Eligible areas and treated and non-treated regions period 1994-2006

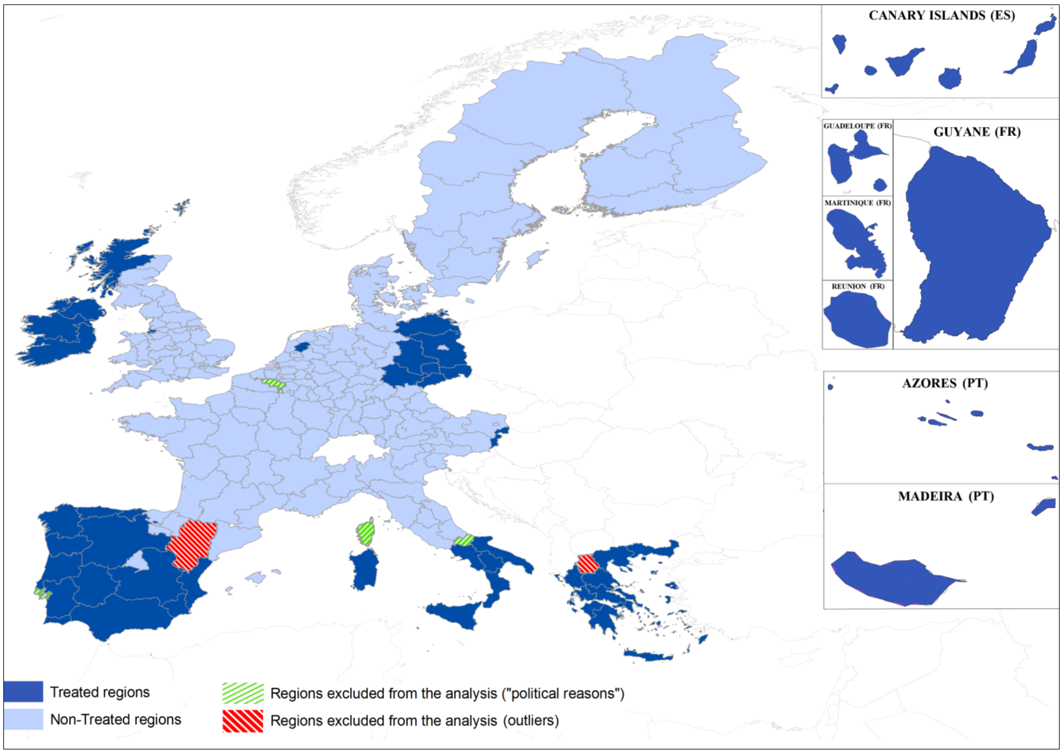
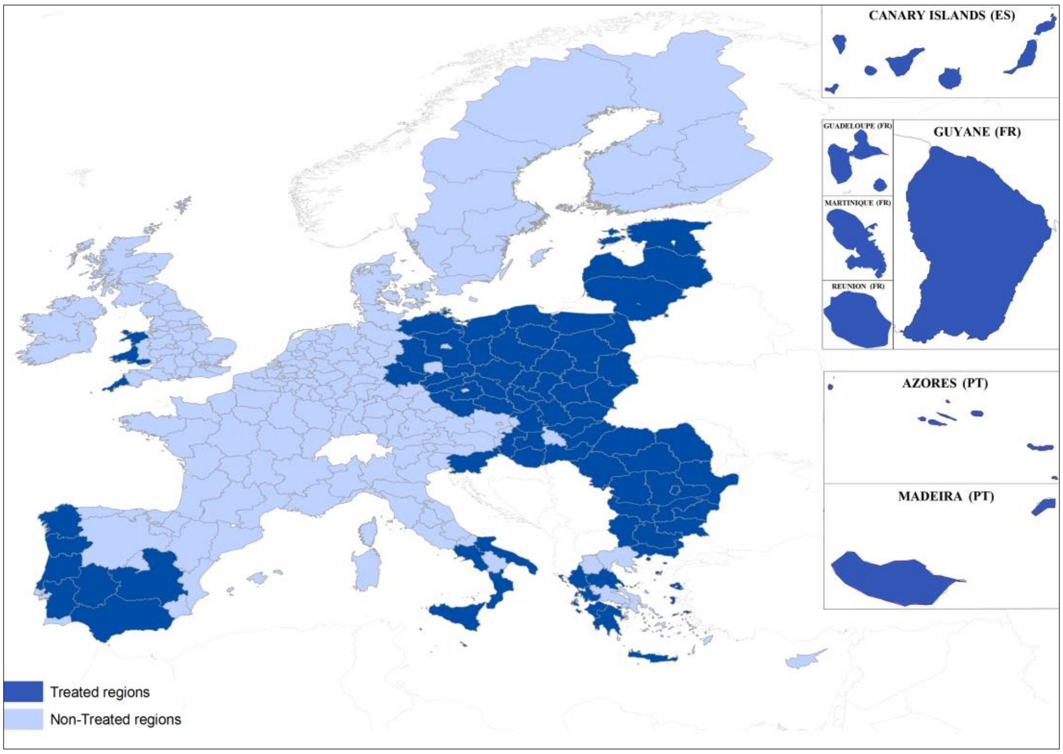


Figure 4 : Eligible areas and treated and non-treated regions period 2007-2013



As the paper is focused on the intensity distribution among European regions, Figure 5 shows the geographical location of the regions with different deciles of treatment intensity (EUF by population) in the EU in the period 1994-2006, Figure 6 in the period 2007-2013. In these figures the core-periphery picture is less clear, indicating that in the determination of the EUF regional intensity several factors were at work. Moreover, the presence of the capping for the new member states regions in the last programming period can explain some differences in intensity among them.

In Table 1 we compare treated and non-treated regions with respect to different covariates in by treatment and programming period. Non-treated regions are generally smaller, but more populated than the treated ones. As expected, they are richer and receive less EUF per capita. However, our data show that the average per capita GDP growth is lower than that of the treated regions. Note that the heterogeneity of the EUF intensity is higher in the first programming period, decreases in the second but increases again in the third one.

3. Econometric Approach

3.1. The impact of continuous treatment in a RDD framework

The key issue when evaluating public policies is to isolate their impact from other factors affecting the outcome under analysis. Our dataset on EUF transfers presents a sharp discontinuity in each period and this allows using a quasi-experimental method deriving from a RD design approach. This enables us to identify the causal effect of transfers on regional growth performances. However, the standard RDD is developed for the case of binary treatment. Our idea, as mentioned in Task 1, is to extend the RDD in the case of continuous treatment, considering the intensity as a cause of the impact heterogeneity. When the treatment is continuous, treatment effects are affected by three components: the treatment level, the heterogeneity among the units and the stochastic component. Apart from the error term, the heterogeneity issue can be interpreted in two ways. First, for each level of treatment, the effects may vary among units: this is the traditional heterogeneity problem in the programme evaluation literature with binary treatment, depending on the characteristics of each unit (covariate heterogeneity). This aspect is relevant for the precision and the unbiasedness of the estimation, but it is not considered in our analysis (see Becker et al., 2013). In this paper we focus on another source of heterogeneity, i.e. the differences in the effects across levels of treatment. This source of variability is handled by evaluating the average effect among units treated at different levels around the discontinuity. Assuming to have an infinite number of observations, a natural development of the treatment effects estimation in the continuous case is the difference between the average outcomes of the units treated at each level with the average outcomes of the untreated units around the cut-off. However, when the number of observations is finite and limited, the heterogeneity in covariate can dominate the heterogeneity in the level of treatment. One alternative is to combine designs and to assume that, after conditioning on covariates, treatment assignment (in differences from the mean for treated and not treated sample) is as-if randomised for those regions near the discontinuity. Therefore, our approach is a combined design, where we consider heterogeneity in RDD after conditioning in pre-treatment covariates. A similar approach, although adopted in a different framework, is presented in Keele et al. (2015). In this paper the problem is the presence of a strong self-selection between small neighbourhoods across treated and not treated areas. Therefore the paper proposes to combine designs and to assume that, after conditioning on the observable variables affecting treatment assignment and the EUF intensity, treatment assignment and the differences from the mean in treatment are as-if randomised around the Ob. 1 assignment threshold.

Figure 5 : Regional distribution of Structural Funds (Intensity = EUF / Population) - Programming periods 1994-2006

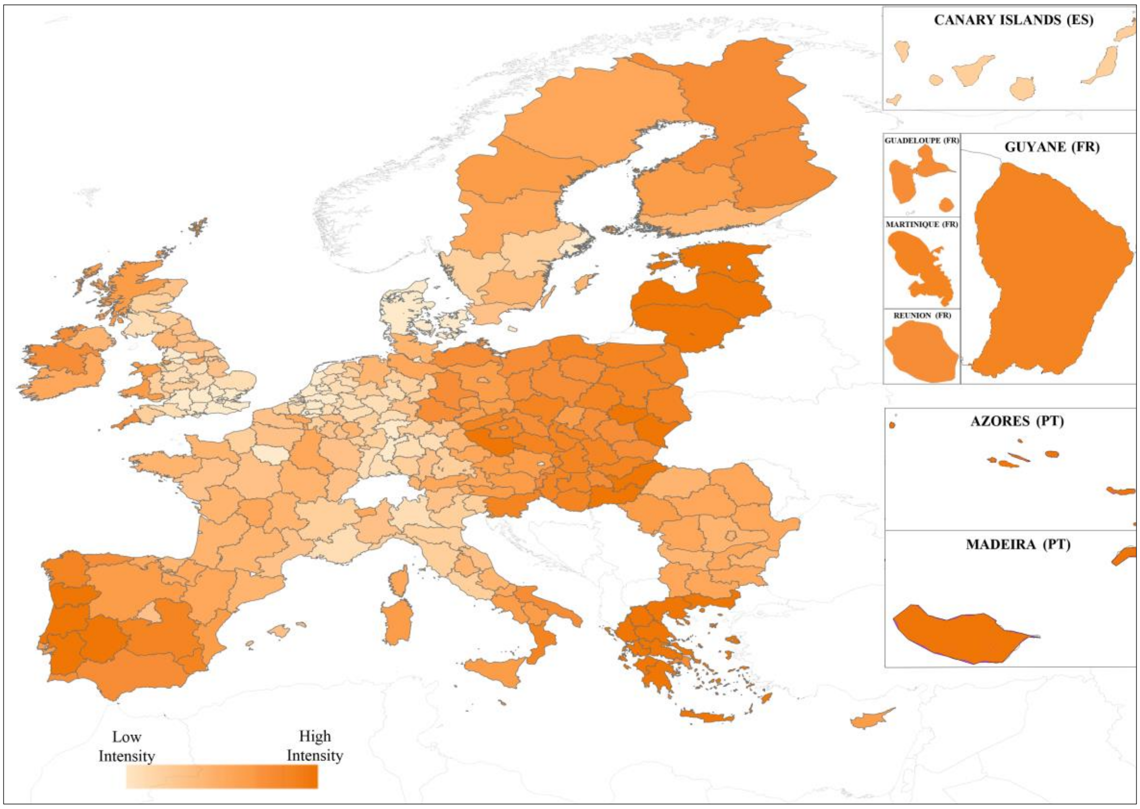


Figure 6 : Regional distribution of Structural Funds (Intensity = EUF / Population) - Programming period 2007-2013

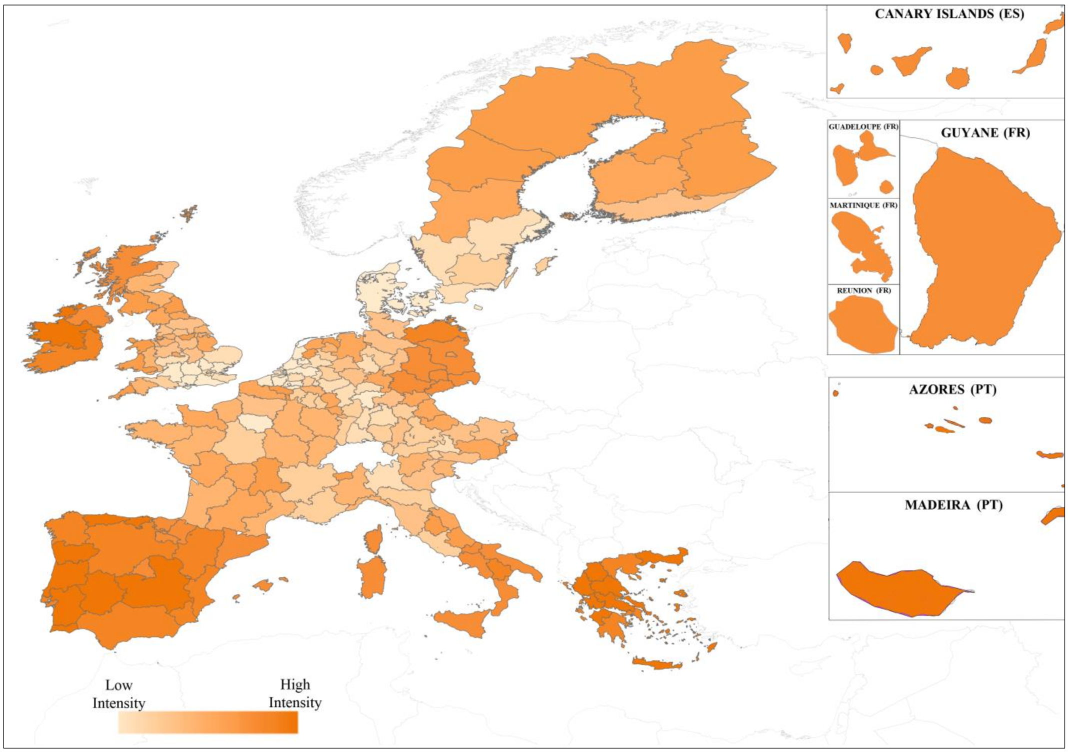


Table 1 : Descriptive statistics (mean) of NUTS2 regions by treatment status and programming period

	1994-1999			2000-2006			2007-2013		
	Treated regions	Non-treated regions	*Total*	Treated regions	Non-treated regions	*Total*	Treated regions	Non-treated regions	*Total*
N° of NUTS-2 Regions	57	150	*207*	53	155	*208*	82	180	*262*
Population *(million inhabitants)*	88	283	*371*	84	294	*378*	155	340	*495*
Total funds *(billion 2005 prices)*	127	42	*169*	111	76	*187*	106	64	*170*
Annual per capita aid intensity (EUR at 2005 prices)	241	24	*76*	190	37	*71*	171	47	*86*
Coefficient of variation of annual per capita aid intensity	0,71	1,32	*1,64*	0,48	0,89	*1,17*	0,63	1,03	*1,05*
Average GDP per capita	14407	24836	*22357*	16318	28916	*26132*	10866	30961	*24660*

Now we formalize our approach, as in Task 1. A simple representation of the EUF framework is the following: we assume two treatment status (S), a status with a high level of treatment (Sh) and a status with a low level of treatment (Sl). In each status the treatment varies in a continuous way around its mean, with the condition E(Sh)>E(Sl). The level of treatment t is defined as the difference from the mean in each status: th= Sh-E(Sh) and tl= Sl-E(Sl). Let us define D=1 if the region is in the status with high level of treatment and D=0 if the region is in the status with a low level of treatment.

The common potential outcome approach in a continuous treatment framework can be applied in our context: yi(T) represents the set of potential outcomes, for each region i, given a random sample indexed by i=1…N and T represents the continuous variable indicating the treatment level, changing from the classical binary definition, named ti .Furthermore, the general observed outcome Y can be written as:

(1) yi = di yi (D=1, ti) + (1-di) yi (D=0, ti)

where D is the dummy variable indicating the treatment status and yi(ti) is the particular potential outcome for each status at the observed level ti. The average treatment effect on the treated at the t-th level is estimated as:

(2) □(T) =E[Y(D=1)-Y(D=0)|T=t]

The parameter □(T) can be defined as the average treatment level effect (ATLE) (see Adorno et al., 2007). However, our analysis is focused on the effect of ti on yi. In a RDD framework, the outcome yi is a function of the treatment di, of the forcing variables xi, and of the level of treatment ti. Our estimate of the ATLE is local, in the sense that it applies in the neighbourhood of the threshold of treatment forcing variable x, for every given ti.

We define the local average treatment level effect (LATLE) at the threshold x0:

(3) $LATLE(xi = x0, ti) = LATLE(x0, ti) = E[y1i |x0, ti] - E[y0i |x0, ti]$

where y1i denotes the outcome with high treatment and y0i the outcome with low treatment, and x0 denotes the threshold values for the forcing variable. The expected value of yi to changes in ti given x=x0 is the dose-response function (DRF) of yi to ti at the threshold:

(4) $yi = DRF(ti|x= x0) = E[yi |x0, ti]$

In our case, the DRF relates each value of the EUF intensity to the GDP growth rate from 1994 to 2006. The estimation of the LATLE and the DRF in a RDD framework requires 3 different identifying assumptions5:

A1. Continuity of outcomes at threshold: E(y1i) and E(y0i) are continuous at x0

This is the standard identifying assumption in the RDD framework: every jump at the threshold must be attributed only to the forcing variable.

A2. Continuity of treatment intensity at the threshold: The variable ti is continuous at x0

Assumption A2 allows identifying the effect of the treatment, based on the average treatment intensity, and the effect of the intensity of the treatment, measured as the difference from the mean, for the treated and the untreated regions. The average jump is attributed to the difference in the average intensity of treatment between treated and not treated regions at threshold.6

A3. Random Assignment of treatment intensity conditional on the forcing variable and the covariates at the threshold: The variable ti is uncorrelated with the error term in the outcome equation, conditional on xi and covariates Zi at the threshold

The assumption states that the treatment intensity (measured as the difference from the mean), conditioned on the forcing variables and other covariates, is randomly distributed between treated and not treated regions. The important condition is that treated and untreated regions having the same level of treatment (in differences from the mean) are not different by some unobservable dimension. The condition is similar to the condition of weak unconfoundedness (CIA) in a GPS framework (Hirano and Imbens, 2004) but it is circumscribed around the threshold.7

3.2. The econometric model

Here we report the general econometric approach adopted, consistent with what we presented in Task 1. In our context, these assumptions state that, conditional on the forcing variable, regions with different levels of treatment around the threshold do not differ in unobserved variables which are relevant for regional GDP growth. Even

---

[5] These assumptions adapt the HLATE framework proposed by Becker et al. (2013) to the case of continuous treatment.

[6] The plot of the treatment intensity distribution in Figure 6 in the next section shows us that assumption A2 is satisfied in our data.

[7] For the use of CIA in a RDD framework see also Angrist and Rokkanen (2016)

around the threshold, in case of a small sample, this condition can require some adjustment for baseline covariates. Therefore, the LATLE is estimated as:

(5) LATLE(x0, ti) = E[y1i |x0, ti,, Zi] – E[y0i |x0, ti, Zi]

where Zi i is a set of baseline covariates. We assume that Zi i captures the characteristics relevant to the probability to receive a relative high or low treatment intensity. Therefore, after controlling for these observable characteristics, any remaining difference in treatment intensity ti across regions is independent of the potential outcome yi.

The same holds for the DRF:8

(6) DRF(ti|x= x0 ,Zi) = E[yi |x=x0, ti,, Zi]

Now we define the parametric control function for identification of the LATLE. We start from the "classic" sharp RDD framework:

(7) GY = a + b0(x) + g*D + b1(x)*D

where GY is the growth rate of per capita GDP, x is the forcing variable (average per capita GDP in different years depending from the programming period) and D is the treatment dummy, equal to 1 when treated, while b0(.) and b1 (.) are sufficiently smooth polynomial functions of x.

Now we assume that the impact g(.) of the treatment is heterogeneous, and depends on t, the relative intensity of treatment (expressed in difference from the mean):

(8) GY = a + b0(x) + g(t)*D + b1(x)*D

Using a polynomial approximation for the term g(t)*D we have:

(9) GY = a + b0(x) + g0(t) + g1D + g2(t)*D + b1(x)*D

where g0(.) and g2 (.) are a sufficiently smooth polynomial functions of t.

In case of a large sample, the heterogeneity would not be a problem for the RDD. However, in our finite sample, we cannot exclude that differences in intensities reflect differences in sample characteristics also around the threshold. As such, we wish to combine identification strategies and assume that, after conditioning on covariates, treatment relative level is locally randomised for those regions close to the threshold. Thus, we propose a mixed design, using both RDD and conditioning on observables (Z):

(10) GY = a + b0(x) + g0(t) + g1D + g2(t)*D + b1(x)*D + h(Z)

Our approach can be explained in two different ways. The first explanation is that we are estimating the intensity effect around the "average treatment impact". Actually we exploit variation in intensity for treated and non-treated regions around the average treatment effect for both groups. If we define the "average or normal effects of treatment given covariates" (GYn), which includes the discontinuity, as:

(11) GYn = a + b0(x) + g1D + b1(x)*D + h(Z)

[8] Essentially, for the correct identification of the DRF assumptions A1 and A3 are sufficient. Besides, it is possible to represent the DRF with respect to the absolute value of the intensity instead of the difference between the intensity and the intensity average of the corresponding treatment group.

where a includes the average intensity effect when the treatment is low, and g1 includes the difference in effect between the average low and high level of treatment.

The conditioned effect of intensity is given by the difference from the "average effect of the treatment given the covariates":

(12) GY - GYn = g0(t) + g2(t)*D

The second explanation is inside the Becker et al. (2013) framework. Intensity can be considered as one of the variables explaining the heterogeneity of the LATE. However, our approach is different in the use of covariates Z: we change Assumption 3 in Becker's paper (Random assignment of the interaction variable conditional on xi), where the interaction variable (which, in our paper, is the relative level of treatment) is uncorrelated with the error term in the outcome equation, conditional on xi (the forcing variable). In our framework the relative level of treatment ti is uncorrelated with the error term conditional on xi and the covariates Z.

4. Results

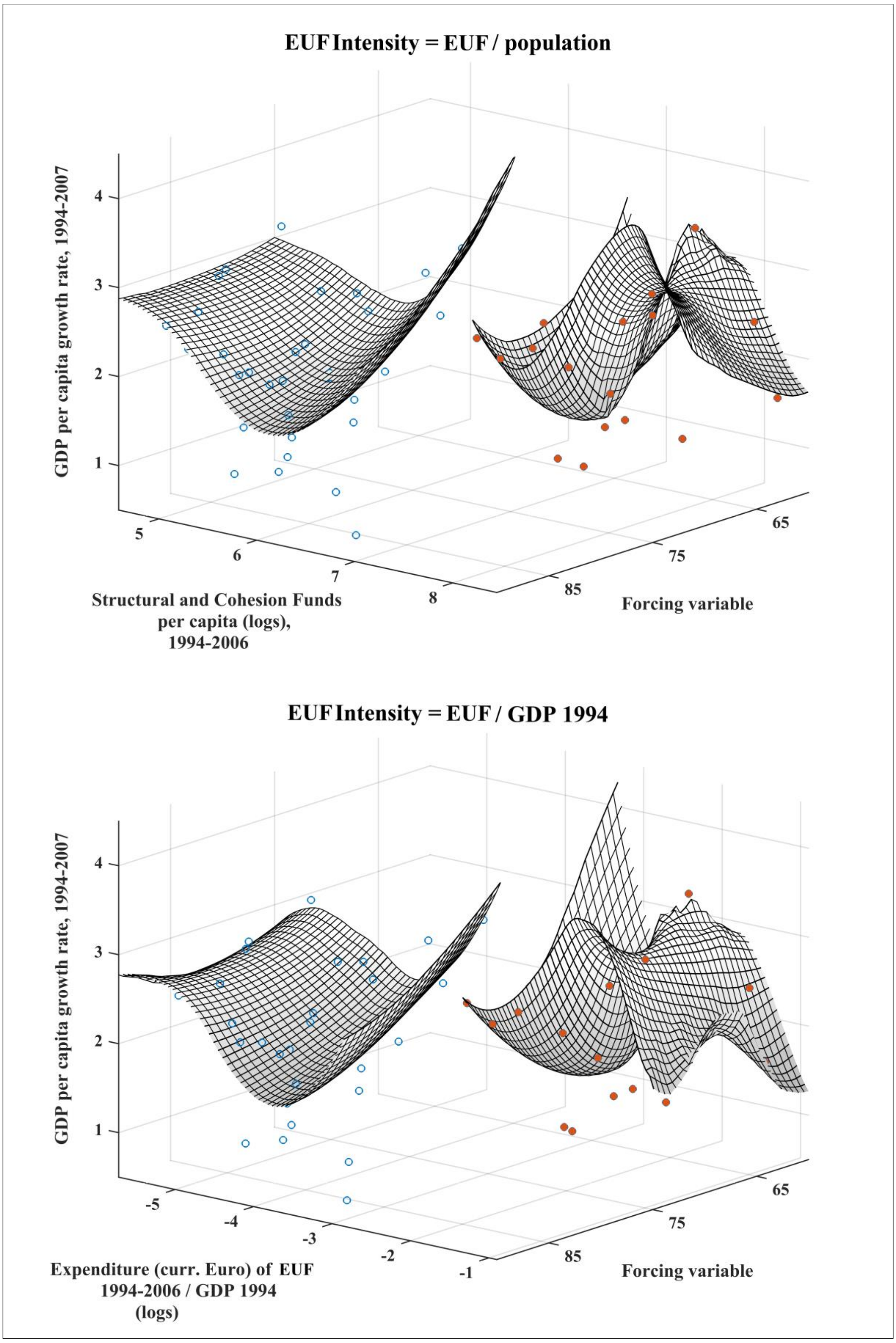
In a RDD analysis it is recommended to represent the relationship between the forcing variable and the outcome with a graph, in order to highlight also visually the presence of a discontinuity (Lee and Lemieux 2010). In our case the problem is more complex, because the three-dimensional graph, showing the relationship between the outcome variable (average annual compound growth rate of *per capita* real GDP), the forcing variable (the level of *per capita* GDP in PPS, EU 15 = 100, the year depends on the programming period) by region, on either sides of the cut-off (75% of EU average GDP per head in PPS), and the intensity, using both the explained normalizations. We show the Figures by different intensity normalization for the period 1994-2006 (Figure 7) and the period 2007-2013 (Figure 8)

In Figures 7 and 8, the cut-off line sharply distinguishes between treated regions (i.e., Ob. 1) and non-treated regions. The surfaces represent quadratic lowess functions (using a bi-square weight function and a bandwidth of 0.8) of the natural log of the forcing variable and the EUF transfers intensity. These functions are estimated on both sides of the threshold separately.

The graph 7 shows the typical shape of the RDD. On average, Ob. 1 regions show higher growth rates than other EU 15 regions, and in the graph this is represented by a clear discontinuity. However, in this paper we are also interested in the relation between intensity and growth. The most interesting aspect of the figure is the concavity that is created in the surface along the intensity axis: the relation between intensity and growth is first steady and then growing among the non-treated regions; while, it increases to a maximum and then decreases for treated regions. This pattern is the same for the two normalisations. The figure then shows how the effect of the intensity on treated regions is not linear, decreasing after the internal maximum.

In graph 8 we note that the difference in the behaviour of treated and not treated regions along the discontinuity is sharper. Moreover, the concavity along the intensity axis is more evident. Therefore the pattern of the period 1994-2006 is confirmed also in the period 2007-2013, even with a reduced time span.

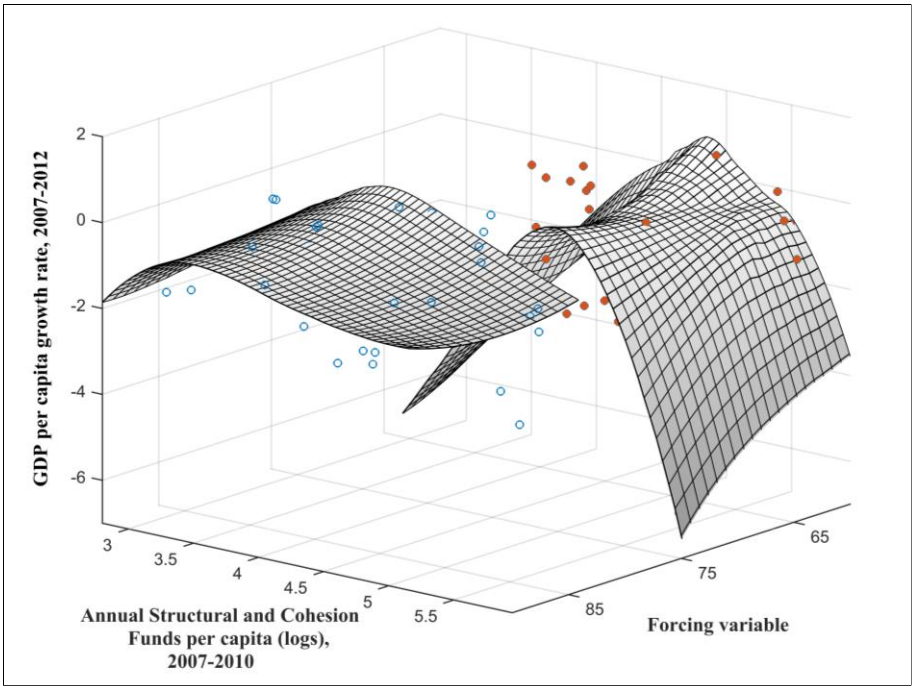
Figure 7: Relationship among the forcing variable, the GDP per capita growth rate and the EUF Intensity (restricted sample) period 1994-2006



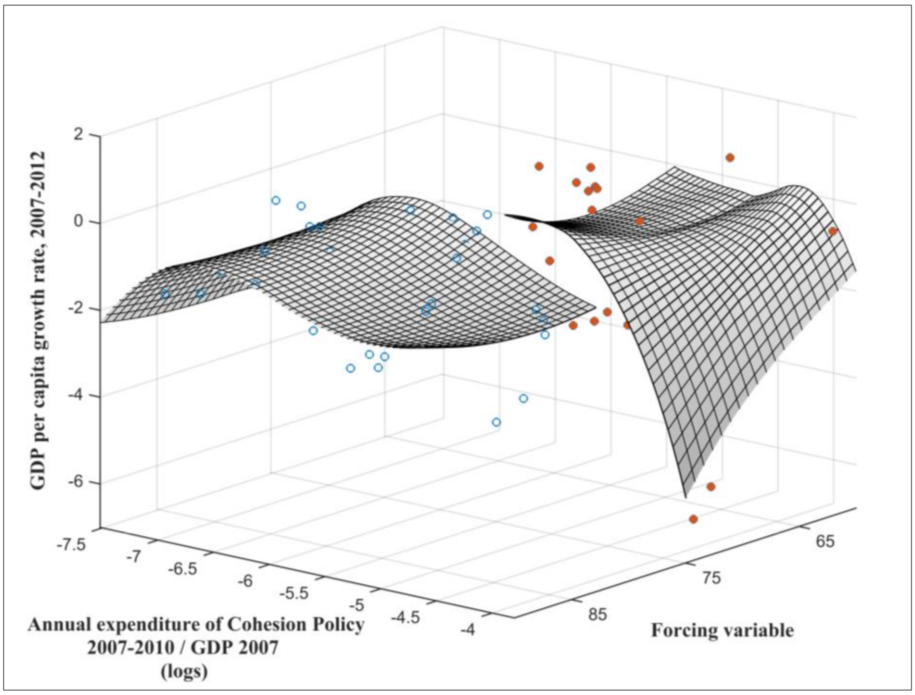
Notes: The upper and lower figures illustrate the relationship between GDP per capita growth rate (1994-2006), forcing variable and EU funds intensity. The solid (hollow) dots indicate regions that were considered (were not considered) Ob. 1 regions. The surfaces represent quadratic lowess functions (using a bi-square weight function and a bandwidth of 0.8) of the forcing variable and subsidy intensity. These functions are estimated on both sides of the threshold separately.

Figure 8 : Relationship among the forcing variable, the GDP per capita growth rate and the EUF Intensity (restricted sample) period 2007-2013

EUF Intensity = EUF / Population



EUF Intensity = EUF / GDP 2007



Notes: The upper and lower figures illustrate the relationship between GDP per capita growth rate (2007-2010), forcing variable and EU funds intensity. The solid (hollow) dots indicate regions that were considered (were not considered) Ob. 1 regions. The surfaces represent quadratic lowess functions (using a bi-square weight function and a bandwidth of 0.8) of the forcing variable and subsidy intensity. These functions are estimated on both sides of the threshold separately.

We present the results of the parametric estimation of the continuous RDD, using the model presented in section 3. With respect to Task 1, the RDD analysis requires a different model specification, that combines the three periods and includes a different number of NUTS-2 regions in each period. Therefore we stack together the three programming periods in a unique dataset, accommodating a different number of NUTS-2 regions in each period. The forcing variables are normalized to 0, in order to match the different samples.

Different order polynomials of the forcing variable can be introduced as regressors in the model, in order to allow different non-linear specifications of the relationship between the outcome and the forcing variable on both sides of the cut-off point. Therefore, the presence of a discontinuity in the relationship between GDP growth and EUF transfers intensity at the threshold cannot be attributed to a missing non-linearity, but exclusively due to the treatment. Lee and Lemieux (2010) argue that, in parametric regressions the choice of the polynomial order is as crucial as that of the bandwidth in the non-parametric approach. After few attempts, we decided to use a third-order polynomial for the forcing variable. We additionally conditioned the equations to some covariates, including surface area, population in 1994, percentage of over 65 in 1994, percentage in the service sector in 1994, hourly productivity in 1994 and employment rate among 15-64 years old in 1994.

Differently to Task 1, we add two different dummies: a dummy for the EU-15 regions, that allows capturing the heterogeneity of EUF effects on new member state regions, and a dummy for each programming period, that allows identifying differences in the effect during the last economic crisis.

The main result of the econometric analysis is that the positive impact of EUF on regional growth is confirmed, also enlarging the analysis to include the estimates of the EUF impacts in the EU-27 regions during the three last programming cycles. However, the statistical significance is not always preserved in all the model specification we estimated. Tables 2 and 3 present the estimates using the intensity expressed as differences from the means, for both definitions of intensity. In these equations, the treatment dummy coefficient and its interactions also capture the effect of the average intensity.

Table 2 : Continuous RDD parametric estimates : deviation from group means - (Intensity = EUF / GDP) – 3 programming periods

Dependent variable: GDP per capita compound growth rate 1994-2001 (for PP 1994-1999), 2000-2007 (for PP 2000-2006), and 2007-2012 (for PP 2007-2013).						
	(1)	(2)	(3)	(4)	(5)	(5 IV)
Dummy Treatment (D)	2.944	2.615	1.950	1.823	1.927	2.187
	(1.322)**	(1.459)*	(1.541)	(1.618)	(1.676)	(1.720)
Intensity	-	-	-0.0031	-0.0011	0.0007	-0.0239
			(0.0020)	(0.0023)	(0.0040)	(0.0202)
Intensity Squared	-	-	-	-0.00002	0.00008	0.0003
				(0.0002)	(0.00011)	(0.0006)
Intensity Cubic	-	-	-	-2.72e-08	-8.81e-07	-4.56e-07
				(8.24e-08)	(9.03e-07)	(5.20e-06)
Intensity*D	-	-	-	-	-0.0029	0.0263
					(0.0046)	(0.0216)
Intensity Squared*D	-	-	-	-	-0.0001	-0.0003
					(0.0001)	(0.0006)
Intensity Cubic*D	-	-	-	-	8.93e-07	3.37e-07
					(9.11e-07)	(5.33e-06)
Dummy EU15	-1.148	-1.173	-1.295	-1.332	-1.292	-1.135
	(1.176)	(1.172)	(1.226)	(1.260)	(1.342)	(1.324)
Dummy PP 00-06	-0.904	-0.898	-0.874	-0.896	-0.892	-0.650
	(0.111)***	(0.111)***	(0.122)***	(0.123)***	(0.144)***	(0.371)*
Dummy PP 07-13	-3.276	-3.272	-3.135	-3.201	-3.219	-2.685
	(0.172)***	(0.173)***	(0.204)***	(0.213)***	(0.248)***	(0.655)***
Dummy EU15*D	-2.239	-1.767	-1.285	-0.742	-0.739	-0.593
	(1.264)*	(1.342)	(1.386)	(1.452)	1.516	(1.789)
Dummy PP 07-13*D	-1.568	-1.550	-1.594	-1.507	-1.504	-1.783
	(0.489)***	(0.499)***	(0.485)***	(0.462)	(0.456)	(0.555)***
Polynomial order forcing variable	1	3	3	3	3	3
Other covariates	si	si	si	si	si	si
Intensity parameters jointly stat. sign. (5% level)	-	-	no	no	no	no

R-squared	0.5827	0.5858	0.5928	0.6022	0.6037	/
Nb. of treated regions	139	139	139	139	139	139
Nb. of non-treated regions	355	355	355	355	355	355

Note: Clustered standard errors at the NUTS2 level in parentheses. The polynomial functions are allowed to have different parameters to the left and the right of the threshold. The estimates are based on the 75% of the sample closest to the Ob. 1 assignment threshold (we exclude the lower quarter for the treated regions and the higher quarter for the non-treated regions). Other pre-treatment covariates include population, population density, percentage of over 65, share in the service sector, productivity and employment rate among 15-64 years old.

***p<0.01, **p<0.05, *p<0.1.

Note that the treatment effect is always positive, and it is statistically significant when we consider the interaction with the third period. The intensity parameters are generally jointly statistical significant at the 5% level when we consider the normalization of intensity with the GDP, but not when the normalization with the population is taken into account. We interact the treatment dummy with the EUF intensity, allowing a different effect of intensity for treated and not treated regions. In this case we have a fully specified model, which is our preferred specification for both normalizations.

Table 3 : Continuous RDD parametric estimates : deviation from group means - (Intensity = EUF / GDP) – 3 programming periods

Dependent variable: GDP per capita compound growth rate 1994-2001 (for PP 1994-1999), 2000-2007 (for PP 2000-2006), and 2007-2012 (for PP 2007-2013).						
	(1)	(2)	(3)	(4)	(5)	(5 IV)
Dummy Treatment (D)	2.944	2.615	2.074	1.994	0.934	0.572
	(1.322)**	(1.459)*	(1.481)	(1.501)	(1.382)	(5.895)
Intensity	-	-	-40.118	-57.169	-43.34	-1,093
			(19.234)**	(25.247)**	(90.23)	(1,259)
Intensity Squared	-	-	-	-4,945	39,929	-49,884
				(2,051)**	(15,565)**	(218,608)
Intensity Cubic	-	-	-	167,211	7,024,965	4.11e+07
				(99,140)*	(2,858,970)**	(3.35e+07)
Intensity*D	-	-	-	-	-19.821	1,163
					(87.717)	(1,282)
Intensity Squared*D	-	-	-	-	34,739	28,377
					(15,527)**	(227,954)
Intensity Cubic*D	-	-	-	-	-6,838,242	-4.11e+07
					(2,854,494)**	(3.31e+07)
Dummy EU15	-1.148	-1.173	-0.912	-0.923	-2.049	-2.602
	(1.176)	(1.172)	(1.205)	(1.179)	(1.059)*	(6.212)
Dummy PP 00-06	-0.904	-0.898	-0.836	-0.875	-0.880	-1.161
	(0.111)***	(0.111)***	(0.125)***	(0.133)***	(0.140)***	(0.280)***
Dummy PP 07-13	-3.276	-3.272	-3.181	-3.198	-3.211	-3.553
	(0.172)***	(0.173)***	(0.184)***	(0.193)***	(0.199)***	(0.400)***
Dummy EU15*D	-2.239	-1.767	-1.553	-0.957	0.165	2.025
	(1.264)*	(1.342)	(1.353)	(1.381)	(1.264)	(5.250)
Dummy PP 07-13*D	-1.568	-1.550	-1.531	-1.550	-1.561	-1.557
	(0.489)***	(0.499)***	(0.482)***	(0.469)***	(0.466)***	(0.668)**
Polynomial order forcing variable	1	3	3	3	3	3
Other covariates	si	si	si	si	si	si
Intensity parameters jointly stat. sign. (5%	-	-	si	si	si	no

level)						
R-squared	0.5827	0.5858	0.5923	0.6065	0.6113	/
Nb. of treated regions	139	139	139	139	139	139
Nb. of non-treated regions	355	355	355	355	355	355

Note: Clustered standard errors at the NUTS2 level in parentheses. The polynomial functions are allowed to have different parameters to the left and the right of the threshold. The estimates are based on the 75% of the sample closest to the Ob. 1 assignment threshold (we exclude the lower quarter for the treated regions and the higher quarter for the non-treated regions). Other pre-treatment covariates include population, population density, percentage of over 65, share in the service sector, productivity and employment rate among 15-64 years old.

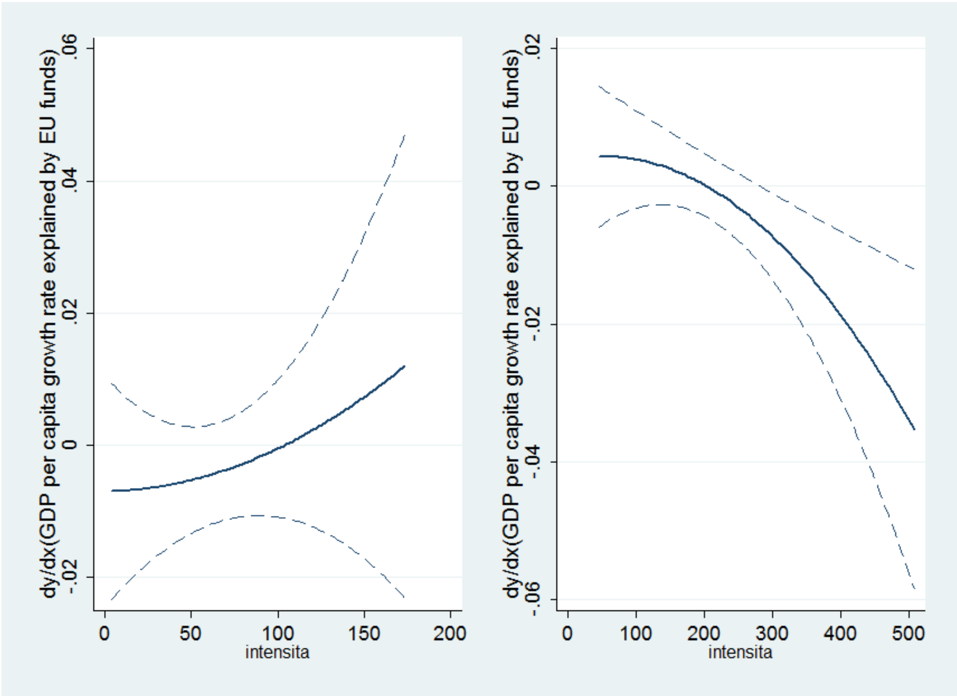
***p<0.01, **p<0.05, *p<0.1.

In the case of the fully specified model and normalized by the population (Table 2), the effect captured by the treatment dummy is high (+1,9 percentage points more in terms of annual GDP growth, when the intensity is on average); this is due to the strong difference in the average treatment between treated and not treated regions. However, this is the effect for the new member state regions. The effect for EU-15 regions is lower, equal to +1.2 percentage points, closer to results presented in Pellegrini et al. (2013), where the average impact is equal to +0.9 in the parametric approach, and the results in Task 1 (+0.7). The effect in the last programming period is lower (+0.4 for the new member state regions, -0.3 for the UE-15 regions). However these effects are not statistically significant in this specification. Only in the IV version the effect of the treatment is positive and statistically significant (see later). Moreover, the intensity parameters are not jointly statistically significant.

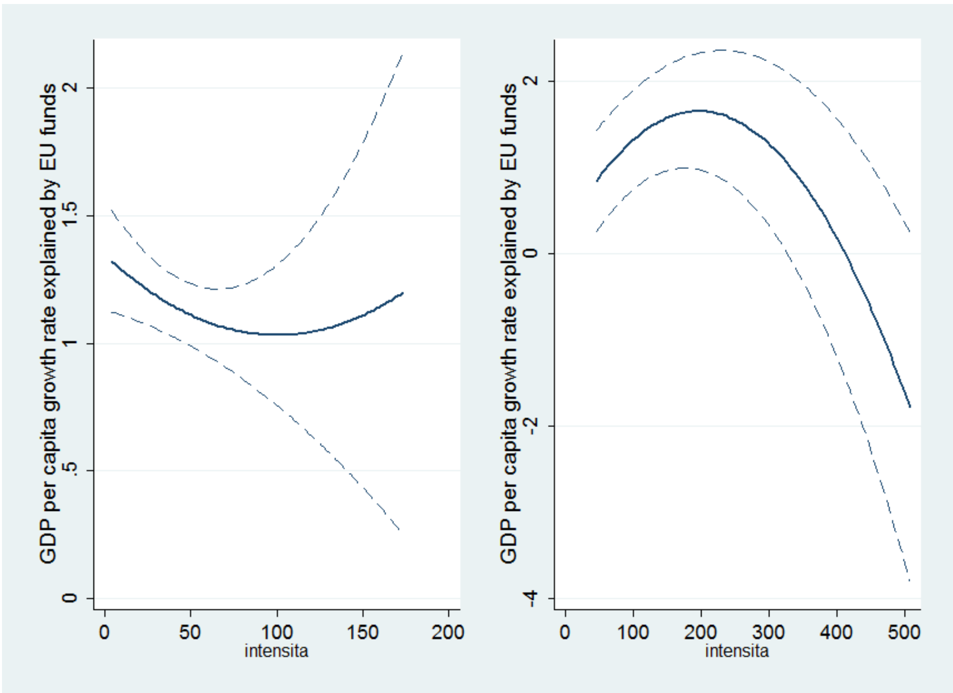
In term of share of GDP, the impact is equal to +0.9 percentage points (ppts) in the new member state regions, +1.1 ppts in the UE-15 regions, like in Task 1. The effect is negative in the last period (-0.6 ppts). In this case, the intensity parameters are jointly statistically significant (Table 3).

Figure 9: The effect of treatment intensity on regional growth (fully specified model) - EUF Intensity = EUF / population

Treatment effect function



Dose-response function



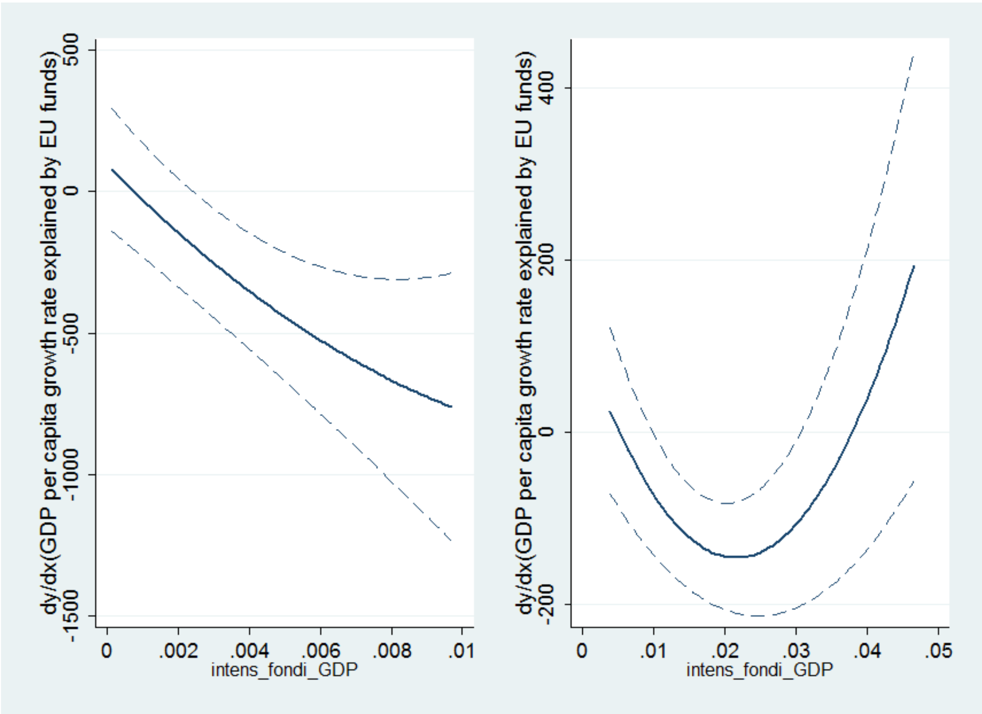
Like in Task 1, also in this sample the relationship between per capita EUF and annual per capita regional growth is decreasing with respect to the EUF transfers intensity. A simple way to represent graphically our results is to draw the curve described by the intensity coefficients of our models.

Using the estimates from the fully specified model (eq. 5 in Table 2), Figure 9 shows the average dose-response function of the GDP per capita growth rate and the EUF transfers intensity and the treatment effect function (the marginal effect of on unit of treatment, i.e. the first partial derivative of DRF) by treatment, both for different level of treatment intensity (EUF by population) and with the 90 % confidence bands. As we can see in Figure 9, the dependent variable is an increasing function of the EUF intensity. The average GDP per capita growth rate is positive for each value of the EUF intensity.

However, the positive impact of the intensity on treated NUTS 2 regions' growth is decreasing and it becomes statistically negligible after a certain threshold. The results using the alternative EUF intensity definition (EUF/GDP, Figure 10) are not so clear, and the dose-response function does not show concavity. However, like Task 1, there is some evidence that NUTS 2 regions receiving lower EUF intensity are much more sensitive to EUF intensity changes than NUTS 2 regions receiving higher EUF intensity levels and that after a certain intensity threshold additional transfers do not increase GDP.

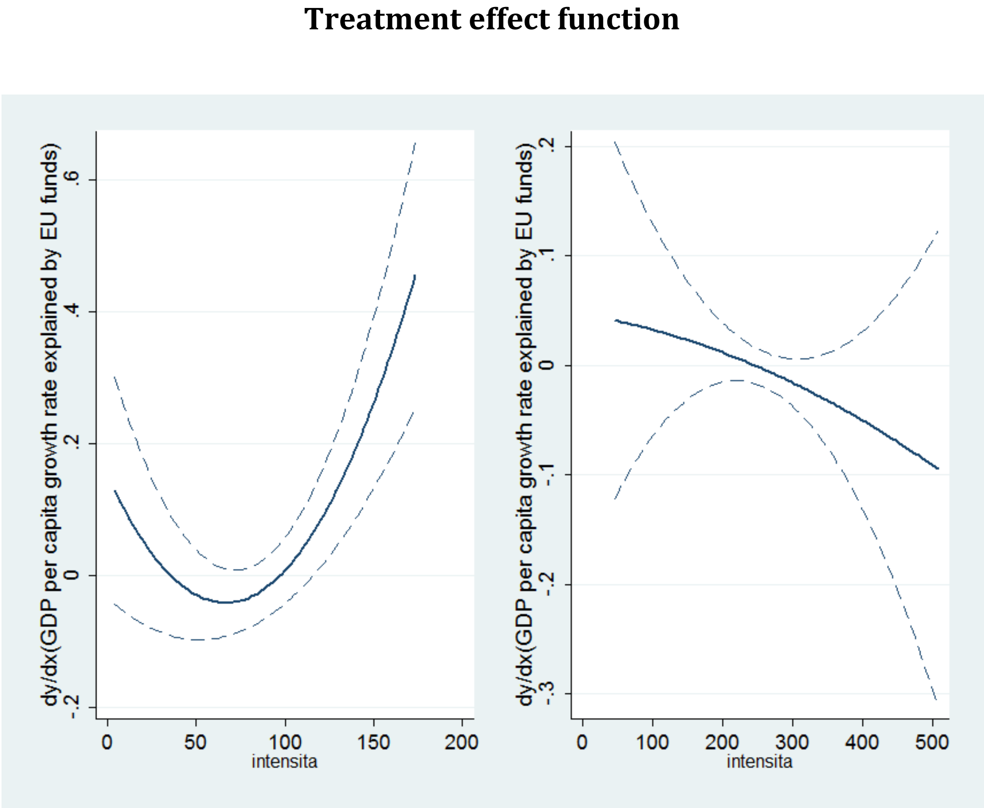
Figure 10 : The effect of treatment intensity on regional growth (fully specified model) - EUF Intensity = EUF / GDP

Treatment effect function



In a standard RDD, the treatment is clearly exogenous to the outcome. However, we cannot exclude that the intensity of the treatment is (partially) endogenous. For instance, regions using efficiently the EUF and growing faster can receive more funds after a middle-period allocation revision. In presence of endogenous treatment intensity our estimates can be biased and the effect of intensity overestimated. Therefore we use an instrumental variables (IV) approach for attenuating this potential issue. As instruments we use a dummy for the cohesion fund countries, the forcing variable relative to the country level, the share of population relative to the country, and the share of employment in the agricultural sector, all covariates estimated in 1994. The results of the IV estimation are presented in the last column of Tables 2 and 3, and draw in Figures 11. The impact of EUF is higher for the normalization by population, slightly lower for the normalization by GDP, and the intensity in not significant in both the normalizations. Figure 11 shows the presence of a decreasing impact also in this specification.

Figure 11 : The effect of treatment intensity on regional growth (fully specified IV model) - EUF Intensity = EUF / population



The main difference with the results of Task 1 is in the enlargement to the programming period 2007-2013, where only data to 2010 are available. The short sample can affect the statistical significance of our estimations. Therefore we replicate the models excluding the last period. The results are presented in Table 4 and Table 5.

In both normalizations the intensity is statistically significant and the results are close to Task 1. Moreover the estimated dose-response functions are concave, showing that the positive impact of the EUF intensity on the growth of the treated regions decreases for more generous regional transfers. Eventually, the impact becomes even statistically negligible after a certain threshold of EUF intensity. Thus, there is additional evidence that the NUTS-2 regions with lower levels of EUF are more sensitive to increases in EUF intensity than the NUTS-2 regions with higher levels of EUF.

Several robustness analyses of our results are common with Task 1, where are already reported, and are not replicated here.

Table 4 : Continuous RDD parametric estimates: deviation from group means - (Intensity = EUF / Population) – 2 different Programming Periods

Dependent variable: GDP per capita compound growth rate 1994-2001 (for PP 1994-1999) and 2000-2007 (for PP 2000-2006)						
	(1)	(2)	(3)	(4)	(5)	(5 IV)
Dummy Treatment (D)	0.209	0.546	0.741	0.716	0.863	1.820
	(0.361)	(0.447)	(0.513)	(0.652)	(0.656)	(1.714)
Intensity	-	-	0.0030	0.0053	-0.0063	-0.0369
			(0.0023)	(0.0025)**	(0.0051)	(0.0380)
Intensity Squared	-	-	-	5.88e-06	0.0003	0.0039
				(0.00001)	(0.0001)**	(0.0020)**
Intensity Cubic	-	-	-	-1.01e-07	-1.10e-06	-0.00003
				(7.04e-08)	(8.96e-07)	(0.00002)*
Intensity*D	-	-	-	-	0.0121	0.0547
					(0.0056)**	(0.0403)
Intensity Squared*D	-	-	-	-	-0.0003	-0.0037
					(0.0001)**	(0.0020)*
Intensity Cubic*D	-	-	-	-	9.94e-07	0.00003
					(8.99e-07)	(0.00002)
Polynomial order forcing variable	1	3	3	3	3	3
Other covariates	1	1	1	1	1	1
Intensity parameters jointly stat. sign. (5% level)	-	-	0	0	1	1
R-squared	0.2109	0.2213	0.2420	0.2668	0.2877	/
Nb. of treated regions	82	82	82	82	82	82
Nb. of non-treated regions	222	222	222	222	222	222

Table 5 : Continuous RDD parametric estimates: deviation from group means - (Intensity = EUF / GDP 1994) – 2 different Programming Periods

Dependent variable: GDP per capita compound growth rate 1994-2001 (for PP 1994-1999) and 2000-2007 (for PP 2000-2006)						
	(1)	(2)	(3)	(4)	(5)	(5 IV)
Dummy Treatment (D)	0.209	0.546	0.750	1.025	1.015	1.584
	(0.361)	(0.447)	(0.523)	(0.745)	(0.778)	(1.093)
Intensity	-	-	27.21	36.37	-164.08	-1,064.06
			(21.34)	(28.31)	(121.64)	(513.30)**
Intensity Squared	-	-	-	-3,202	115,910	849,790
				(2,087)	(50,960)**	(328,839)**
Intensity Cubic	-	-	-	29,195	-8,696,203	-1.04e+08
				(105,758)	(5,833,448)	(4.72e+07)**
Intensity*D	-	-	-	-	183.73	1,038.03
					(118.50)	(487.53)**
Intensity Squared*D	-	-	-	-	-119,918	-861,583
					(50,637)**	(328,348)***
Intensity Cubic*D	-	-	-	-	8,773,640	1.05e+08
					(5,832,965)	(4.72e+07)**
Polynomial order forcing variable	1	3	3	3	3	3
Other covariates	1	1	1	1	1	1
Intensity parameters jointly stat. sign. (5% level)	-	-	0	0	1	1
R-squared	0.2109	0.2213	0.2300	0.2633	0.2917	/
Nb. of treated regions	82	82	82	82	82	82
Nb. of non-treated regions	222	222	222	222	222	222

Note: Clustered standard errors at the NUTS2 level in parentheses. The estimates are based on the 75% of the sample closest to the Ob. 1 assignment threshold (we exclude the lower quarter for the treated regions and the higher quarter for the non-treated regions). Other covariates include population in 1994, population density in 1994, percentage of over 65 in 1994, percentage in the service sector in 1994, productivity in 1994 and employment rate among 15-64 years old in 1994. Note: ***p<0.01, **p<0.05, *p<0.1.

5. CONCLUSIONS

The intensity of EUF transfers is highly heterogeneous across regions, even within the same country. This paper focuses on both the average impact of EUF (Structural and Cohesion funds) and the heterogeneity of the treatment intensity. We use a regional dataset, which is fully coherent with Structural Funds Regulations, three programming periods from 1994 to 2013, and the spatial grid defined by the EU27 regions at level 2 of the 2006 NUTS classification. We propose a new method for estimating the effects of intensity on growth, extending the RDD framework to the case of continuous treatment.

The results basically confirm the positive effects of EUF transfers on regional growth presented in Task1. The findings of this evaluation show that regional policies have a role for stimulating regional growth and economic development. Even considering transfer heterogeneity, the results show a positive effect of EUF on regional growth. The effects are in line with the results of other studies casted in a counterfactual framework.

However, the results are less statistically significant than in Task 1. The main reason is the limited availability of data. We need additional information (beyond the year 2010) for a more robust empirical analysis of the last programming period (2007-2013), where the heterogeneity across regions is higher, due to the presence of new Member States and the largest economic crisis in Europe since WWII was in action. Our results show a higher impact of EUF for the new member state regions, and a lower impact during the crisis. However, the empirical findings for the last programming period will have to be confirmed when the complete data will be available.

Another important result is that the estimated marginal impact on regional growth of further increasing the intensity of the EUF tends to be higher on average for the regions that do not already receive a high intensity of the EUF also in UE-27 regions. In other words, the marginal impact on growth of adding more EUF intensity tends to decrease for the regions with high EUF intensity. A great deal of caution, however, should be exerted in interpreting these results as supportive of the hypothesis that diminishing returns to investment and/or limited absorption capacities may be in place to hamper the full economic-development potential of the high intensities of the EUF transfers.

However, GDP is only one dimension on the target of EU regional policy, which is oriented to remove economic and social disparity across European regions. Therefore effects on GDP are important but cannot exhaust the purpose of the regional policy. Moreover, the decreasing of the marginal returns of transfers can be differentiated by target. For instance, the impact could be always positive in term of employment or social inequality. Therefore this analysis can give some suggestion on interregional allocation of EUF, but cannot be the base of a new allocation of EUF transfers among European regions.

Moreover, the strong limitations in the current data-availability scenario on the EUF payments do not enable to further investigate other important conditions under which the different intensities of the EUF produce desirable regional growth outcomes. These conditions are, for example, the different compositions and scopes of the actual programme interventions, the duration of the project implementations (which may affect the temporal lag needed to observe the desirable results in terms of regional growth), the intensities of the national or regional sources of public aids that may affect the regional growth outcomes in conjunction with the EUF. Thus, also in our case, the empirical analysis cannot aim at offering conclusive evidence on the exact conditions under which the different intensities of the EUF produce desirable regional growth outcomes.

6. References

Adorno V., Bernini C., and Pellegrini G. (2007). "The impact of capital subsidies: new estimations under continuous treatment", Giornale degli Economisti e annali di economia, 120 (66): 67-92.

Angrist J. D., and Rokkanen M. (2016). "Wanna get away? Regression discontinuity estimation of exam school effects away from the cutoff", Journal of the American Statistical Association, 110(512): 1331-1344.

Anselin L. (2006). "Spatial econometrics". In: T. Mills, K. Patterson, eds. Palgrave Handbook of econometrics, vol. 1. In: Econometric Theory. Basingstoke: Palgrave Macmillan, 901-969

Barbieri G., and Pellegrini G. (1999). "Coesione o sgretolamento? Un'analisi critica dei nuovi fondi strutturali", Rivista economica del Mezzogiorno, a. XIII n. 1-2.

Becker S.O., Egger P.H., and von Ehrlich M. (2012). "Too much of a good thing? On the growth effects of the EU's regional policy", European Economic Review, 56 (4): 648-668.

Becker S.O., Egger P.H., and von Ehrlich M. (2013). "Absorptive capacity and the growth and investment effects of regional transfers: a regression discontinuity design with heterogeneous treatment effects", American Economic Journal: Economic Policy, 5 (4): 29-77

Becker S.O., Egger P.H., and von Ehrlich M. (2016). "Effects of EU Regional Policy: 1989-2013", Warwick Economics Research Paper Series, March, 2016, Series Number: 1118 .

Behrman J., Cheng Y., and Todd P. (2004). "Evaluating preschool programs when length of exposure to the program varies: a nonparametric approach", The Review of Economics and Statistics, 86 (1): 108-132.

Cattaneo M.D. (2010). "Efficient semiparametric estimation of multi-valued treatment effects under ignorability", Journal of Econometrics, 155: 138-154.

Drukker D.M., Prucha I.R., and Raciborski R. (2013). "Maximum likelihood and generalized spatial two-stage least-squares estimators for a spatial-autoregressive model with spatial-autoregressive disturbances", Stata Journal, 13: 221-241.

Hahn J., Todd P., and van der Klaauw W. (2001), "Identification and estimation of treatment effects with a regression-discontinuity design", Econometrica, 69 (1): 201-209.

Hirano K., Imbens G.W. (2004). "The propensity score with continuous treatments". In: Andrew G., Meng X.-L. (Eds.), Applied Bayesian modeling and causal inference from incomplete-data perspectives, Wiley, 73-84.

Imai K., Van Dijk D.A. (2004). "Causal inference with general treatment regimes: generalizing the propensity score", Journal of the American Statistical Association, 99: 854-866.

Keele L., Titiunik R., and Zubizarreta J. (2015). "Enhancing a geographic regression discontinuity design through matching to estimate the effect of ballot initiatives on voter turnout", Journal of the Royal Statistical Society: Series A, 178 (1): 223-239.

Lee D.S., and Lemieux T. (2010). "Regression discontinuity designs in economics", Journal of Economic Literature, 48 (2): 281-355.

Mohl P., and Hagen T. (2010). “Do EU structural funds promote regional growth? New evidence from various panel data approaches”, Regional Science and Urban Economics, 40(5): 353-365.

Muccigrosso T. (2009). “La valutazione degli effetti delle politiche di coesione dell’Unione Europea sulla crescita regionale”, unpublished Ph.D. Tesis, University of Bologna.

Pellegrini G., Terribile F., Tarola O., Muccigrosso T., and 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.

Rodríguez-Pose A., and Garcilazo E. (2015). “Quality of government and the returns of investment: examining the impact of cohesion expenditure in European regions”, Regional Studies, 49 (8): 1274-1290.

Thistlethwaite D.L., and Campbell D.T. (1960). “Regression-discontinuity analysis: an alternative to the ex post facto experiment”, Journal of Educational Psychology, 51 (6): 309-317.

HOW TO OBTAIN EU PUBLICATIONS

Free publications:

- one copy:
  via EU Bookshop (http://bookshop.europa.eu);

- more than one copy or posters/maps:
  from the European Union’s representations (http://ec.europa.eu/represent_en.htm);
  from the delegations in non-EU countries
  (http://eeas.europa.eu/delegations/index_en.htm);
  by contacting the Europe Direct service (http://europa.eu/europedirect/index_en.htm)
  or calling 00 800 6 7 8 9 10 11 (freephone number from anywhere in the EU) (*).

  (*) The information given is free, as are most calls (though some operators, phone boxes or hotels may charge you).

Priced publications:

- via EU Bookshop (http://bookshop.europa.eu).

Priced subscriptions:

- via one of the sales agents of the Publications Office of the European Union
  (http://publications.europa.eu/others/agents/index_en.htm).

KN-02-16-718-EN-N

Publications Office

doi: 10.2776/3077