Buying Support for Unification: The Impact of Cohesion Policy on Public Opinion towards EU integration

Kristoffer Johansson*

Master's Thesis in Policy Economics

Erasmus School of Economics

September 2017

* Student number: 448341kj. Email: 448341kj@eur.nl. Supervisor: Dr. Olivier Marie

Abstract

This thesis argues that the funds and financial instruments of the EU Cohesion Policy are important facilitators of public opinions favorable to the EU. By using a fuzzy regression discontinuity design, the results of this thesis suggest a positive effect of the EU Cohesion Policy funds on public support for EU integration and unification with regions receiving funds under the convergence objective being on average 8-15 percentage points more in favor of EU integration and unification than the control regions. This can be explained by the increased regional GDP per capita and real income level previous studies found to be the effect of the Cohesion Policy. The result is in line with literature, particularly with Massetti & Schakel that found a similar but larger effect on the funding's effect on regional party positions towards the EU. The results are sufficiently robust to a number of different specifications and bandwidths, however the number of observations is quite low due to data availability hence one should be careful to draw any strong conclusions based on this thesis. Nevertheless, this thesis adds to the importance of the Cohesion Policy as an instrument to keep the union together and further European integration.

Keywords: European Union, Cohesion Policy, Regional funding, Euroscepticism, Public Opinion

Table of Contents

List of Tables

List of Graphs

List of Equations

Introduction

The European Union is in trouble and has been for several years now. With Brexit and other countries' waning support for the European project, the future of the European Union is uncertain. The gloomy economic mood with many regions experiencing high unemployment and public debt as well as low growth since the financial crisis has contributed to the rise of more Eurosceptic parties and opinions (Nicoli, 2016) and has reduced public faith in both the EU and national governments (Armingeon & Ceka, 2013). An increased perceived threat from immigration, from both outside and within the EU, contributed to this recent rise in Euroscepticism (Treib, 2014). One may wonder what the EU can do to keep the union intact and improve their popular opinion. This thesis suggests that the funds and financial instruments of the EU Cohesion Policy are important facilitators of public opinions favorable to EU integration and unification.

More than one-third of the EU budget is currently allocated to the financial instruments of EU Cohesion policy (European Commission, 2014). The Structural Funds and the Cohesion Fund are financial tools set up to implement the regional policy of the European Union. They aim to reduce regional disparities in income, wealth and opportunities. Europe's poorer regions receive most of the support, but all European regions are eligible for funding under the policy's various funds and programs. The funds are distributed between different objectives of which by far the largest is the Convergence Objective (formerly known as Objective 1). For a region to be eligible for funding under this objective the region needs to have a GDP per capita below 75% of the EU average. A large volume of research has been published on the economic effects of the Cohesion Policy which has shown a moderate positive effect on growth and income among recipient regions (Pieńkowski & Berkowitz, 2015). Economic literature suggests that economic benefits are important drivers of further EU unification and integration, and that a large difference in GDP per capita or income distribution between regions can make European integration very difficult. Large differences in GDP per capita and income levels between regions implies lower economic benefits than costs of unification (Bolton & Roland, 1997) and also creates incentives for migration from poorer to richer regions which in turn might result in a more Eurosceptic population if they perceive immigration as a threat (Hooghe & Marks, 2005; Galgóczi, Leschke, & Watt, 2009).

As the core objective of the regional funding of the EU Cohesion Policy is to reduce these regional disparities, we should expect that these funds increase the public support of EU integration and unification in recipient regions and that, all else equal, regions receiving larger amounts of funds will have higher public support than those who receive less. This hypothesis is not rejected and the result of this thesis suggests that the public support of EU integration and unification is between 8-15 percentage points higher for regions receiving convergence funding compared to the control regions.

This thesis is structured as follows; Section 1 covers theory, previous research and literature. Section 2 presents the research design along with the data, and Section 3 assesses and validates the internal validity of the research design. Section 4 contains the baseline results and Section 5 evaluates the robustness of the results. Finally, Section 6 concludes.

1. Literature Review

A. Theorizing public support for EU integration

In this subsection, I will discuss two different theories that help explain the factors behind public support for EU integration. First, I will account for the economic theory behind unification as developed by Bolton & Roland (1997). Second, I will present social identity theory which helps us explain non-economic factors that influence support for EU integration, e.g. migration.

Economic theory for unification

The driving factors of EU unification has long been disputed but there is a wide consensus that economic factors play a very large role. In an influential paper by Bolton & Roland (1997) they develop a model for the breakup or unification of nations that is especially relevant to, and largely focused on, European countries and the EU. Their analysis is focused on some important economic and political determinants of the process of unification and separation. Their analysis takes its starting point at an economic efficiency point of view where separation of nations is never desirable. A unified nation is always more efficient since free trade among regions is guaranteed, and several public goods are easier coordinated. However, the benefits of unification cannot in general be evenly distributed among all citizens. In each region, there may be winners as well as losers from regional independence. In a democratic context, the question is then whether there is a majority of winners supporting separation, regional autonomy, or unification.

Bolton & Roland focus on regional conflicts over fiscal policy arising from differences in income distribution across regions. In a union, the equilibrium tax rate will generally not coincide with the tax rate chosen by a majority in each region since the income distribution in each region is not identical and regions within the union do not have total freedom in their choice of tax policies. Separation removes any institutional constraints imposed by the union and allows for policies that are closer to the wishes of a majority of voters in the region. When contemplating a move towards independence, voters in each region must then weigh the efficiency benefits of the union against the benefit of having a redistribution policy closer to the preferences of the region.

More recent literature has taken a similar approach with for example Besley & Coate (2003) focusing on differences in tastes for public spending as an intra-union conflict that might induce separation, and Lockwood (2002) concentrating on potential efficiency gains in public spending from separation.

In Bolton & Roland there are basically three different factors influencing a region's choice to separate (1) a political factor that arises from differences in regional preferences over fiscal policy; (2) the efficiency losses from separation; (3) a tax base factor that emerges whenever per capita income varies across regions.

A key insight of Bolton & Roland is that European unification is facilitated by reducing both differences in per capita income across member states and regions, and differences in income distribution. Thus, as these are the core objectives of the regional funding of the EU Cohesion Policy we should expect the funds to have a positive effect on public support for EU unification and integration.

Social identity theory

Of course, not all drivers of unification are economic. Social identity theory, originally developed by Turner, Brown, & Tajfel (1979), proposes that individuals have the fundamental need to perceive their group as superior to many other groups and subsequently apply favorable characteristics to themselves those they perceive as members of their group via a mental process labeled 'social identification', and they value other groups negatively via mechanisms of 'social contra-identification'. Hence, to the extent that we may assume that in-group bias translates into attributing higher value to (national) traditions, we may infer that these national sentiments drive resistance to policies directed toward 'integrated nation-states', which consequently may be considered to be at the expense of national sovereignty (Hooghe & Marks, 2005). Similarly, immigration and the presence of foreigners, legitimized by EU regulations, may be considered a threat to national traditions, irrespective of whether the foreigners are from within the EU.

Lubbers & Scheepers (2007) investigates different factors related to immigration, political trust, income and education levels to explain how they are connected to public support for EU unification and integration. Their main conclusion is that perceived threat from immigrants as well as political distrust increase political Euroscepticism especially among lower-skilled workers, which in turn explain lower levels of Euroscepticism among higher educated people and higher income categories.

The two main drivers of intra-EU migration are the access to labor market and effective increase in earnings potential (Galgóczi, Leschke, & Watt, 2009). The income inequalities within the EU has caused migration from less developed to more developed members, made possible by Schengen (Galgóczi, Leschke, & Watt, 2009). The threat perceived from immigration played a big role in the recent Brexit vote (Wincott, Peterson, & Convery, 2017) as many Britons fear for both their jobs and more importantly, their national identity.

B. Empirical research on Cohesion Policy

A large volume of literature has been published on the effects of EU Cohesion Policy. Pieńkowski & Berkowitz (2015) compiled a summary report published by the European Commission on econometric assessments of Cohesion Policy growth effects containing over 20 different studies by different authors. One should of course be careful when drawing conclusions from this report since the EU has incentives to publish results in line with their agenda. With this said, the vast majority of the studies find a modest positive effect on growth although many suggest that the funding has not been allocated as efficiently as it could have and that gaps between regions are not reduced as quickly as they could if the funds came with certain conditions on sound fiscal policies.

Becker & Eggert (2010) and Pellegrini et.al. (2013) both evaluate the effect of the funding on growth in a regression discontinuity framework and both find a positive effect of 0,6-1,6 percentage points additional growth for treated regions eligible under the convergence objective compared to the control group. Their research designs are valid and results robust, although Pellegrini et.al can be criticized for a non-random exclusion of a few outlier regions that receive far more funding per capita than comparable regions eligible under the same funding objective. Research has also shown a positive effect of the funding on real income levels in treated regions (Ramajo, Márquez, Hewings, & Salinas, 2008) which in turn should result in greater public support of the EU (Gabel & Whitten, 1997) and benefits of redistribution has been an important influence on countries' decision to whether be a part of the EU or not (Doyle & Fidrmuc, 2006). The empirical findings of higher relative growth and income levels as a result of the EU Cohesion Policy funding should theoretically increase the public support for EU unification and integration (Bolton & Roland, 1997) and the closing of the income gap between less and more developed EU members should decrease incentives for intra-EU migration (Galgóczi, Leschke, & Watt, 2009) making immigration a smaller concern when debating EU integration and unification.

Although most studies have focused on the economic effects, Massetti & Schakel (2016) studied the impact of Cohesion funds on regionalist parties' position on European integration. They studied three different funding periods and concluded that the funds had a positive effect on regionalist parties' level of support of European integration and also suggest that the funding has been an important facilitator for the spreading of Europhile positions as well as represented a sort of barrier against the diffusion of Eurosceptic positions. Massetti & Schakel took an ordered logit approach suitable to their research on party positions. Their results are robust to different specifications; however, the authors give little information about the internal validity of their identification strategy and how they address selection bias.

In this thesis, I will take a regression discontinuity (RD) approach, an approach if properly executed is a very strong identification strategy able to find causal effects. I hope to answer whether the same effect Massetti & Schakel found on regionalist parties holds for the general public when using RD, that is, whether the funding has a positive effect on public support for EU integration and unification.

Based on above theory and literature, I form the following hypothesis;

H_0: The regional funding of the EU Cohesion Policy increases the public support of EU integration and unification in recipient regions and, all else equal, regions receiving larger amounts of funds have higher public support than those who receive less.

2. Research Design

A. *Eligibility for funding*

During the 2007-2013 funding period, the European Regional Development Fund, Cohesion Fund and European Social Fund were used for three objectives: (1) Convergence, (2) Regional Competitiveness and Employment and (3) European Territorial Cooperation. The proportion allocated to the Convergence objective was and still is by far the largest and amounted to €282.8 billion, representing 81.5% of the total. The funds were allocated on a NUTS2-2003 regional level where regions with a GDP per capita below 75% of the EU-27 average are eligible for the 'Convergence' objective and all other regions have access to funds under the 'Regional Competitiveness and Employment' objective and the 'European Territorial Cooperation' objective. However, in addition to those eligible for Convergence under the 75% threshold criteria there are a few regions receiving convergence funding on a phasing-out basis. These regions used to be under the threshold but due to the statistical effect of EU enlargement now have a GDP per capita slightly above it. (European Commission, 2008)

The European Commission used regional 2001-2003 (ESA95) data estimated by Eurostat to establish the list of regions eligible for funding under the Convergence objective for the 2007-2013 funding period (Eurostat, 2010).

Data on regional economic accounts (ESA95) is available at Eurostat regional statistics under NUTS2 classification (Eurostat, 2017) and the European Commission InfoRegio provides estimations on funding allocations and expenditures broken down by NUTS2 region and objective (Inforegio, 2017).

B. *Evaluation Design*

The process in which selection for convergence funding was determined allows us to evaluate the causal impact of the EU Cohesion Policy on public views on EU integration using a regression discontinuity (RD) design. Due to the nonperfect compliance, caused by the phasing-out regions, a fuzzy design should be used. Fuzzy RD exploits discontinuities in the probability of treatment conditional on a covariate. This results in a research design where the discontinuity becomes an instrumental variable for treatment status. This naturally leads to a simple 2SLS estimation strategy (Angrist & Pischke 2009).

Fuzzy RD is a vastly superior identification strategy compared to simple OLS. If the effect of the funding would be investigated using an OLS-regression the result would suffer from selection bias which subsequently means that the regression will not yield the causal estimate. This is due to the fact that a simple comparison of support of EU integration between regions that did and did not receive convergence funding fails to account for unobserved factors that may be correlated to both whether a region receives the funds and what its opinion towards EU integration is. As a result, the treated regions might be inherently different than the non-treated regions, and a comparison would make us erroneously attribute the differences in support for integration between the treated and non-treated regions to the funding when it is in fact due to unobserved differences. Regression discontinuity overcomes this selection bias because regions just below the assignment threshold, 75% of GDP per capita of the EU average, are likely very similar to regions just above it, hence treated and non-treated regions are comparable just around the threshold. This means that effect of the funding will be consistently estimated if the regional GDP per capita of the EU average is the only systematic determinant of whether a region receives the funding or not. If this is the case, the funding will not be correlated to unobservable factors and the error term, hence it is possible to estimate the causal effect.

In this thesis, identification of the causal effects is achieved though the inclusion in the first stage regression of a dummy variable that records the exogenous change in eligibility for convergence funding that happens at the threshold. The dummy variable ELIGIBLE is assigned to 1 if a region has below 75% in GDP per capita of the EU average. The estimated first stage equation relates the likelihood of receiving treatment, i.e. convergence funding, to being eligible for treatment:

$$Treatment_{jk} = \beta_0 + \beta_1 ELIGIBLE_{jk} + f\left(GDPofavg_{jk}\right) + \gamma Z + \varepsilon_{ijk} \quad (1)$$

The reduced form is:

$$Unif.Support_{ijk} = \delta_0 + \delta_1 ELIGIBLE_{jk} + f\left(GDPofavg_{jk}\right) + \delta X_{ijk} + \varphi Z + \varepsilon_{ijk} \quad (2)$$

The structural equation therefore used to find causal estimates is:

$$Unif.Support_{ijk} = \theta_0 + \theta_1 Treatment_{ijk} + f\left(GDPofavg_{jk}\right) + \vartheta X_{ijk} + \sigma Z + \varepsilon_{ijk} \quad (3)$$

The treatment variable $Treatment_{jk}$ indicates whether a region receives the convergence funding or not. $Unif.Support_{ijk}$ represents the outcome in support for EU integration and unification. This variable is based on an ESS survey question about whether the European unification process should go further or has gone too far, with responses on a scale from 0 to 10, with 0 being the lowest support for unification and 10 the highest. For simplicity, we can of 0 as 0 % support of EU integration and 10 as 100% support.

The subscript i indicates the individual in region j in country k. The control variable X_{ijk} is a vector of individual characteristics, Z is a vector of fixed effects and ε_{ijk} are the error terms which are clustered at the level of treatment, i.e. NUTS2 regions. Lastly, $f(GDPofavg_{jk})$ is a polynomial expansion of the assignment variable '$GDPofavg$' which represents the regional GDP per capita as percentage of EU average.

The reason we include a polynomial function of the assignment variable is that failure to do so could potentially induce a bias into the estimation of the treatment effect, since this estimate relies on local estimation. While a linear approximation of a non-linear process can be argued to minimize specification errors globally, it nevertheless allows for large specification errors at a specific point such as the cutoff value (Lee & Lemieux, 2010). Imposing a linear structure for a non-linear process might lead us to inappropriately attribute some effect to a discontinuity that actually just represents non-linearity (Angrist & Pischke, 2009).

The fixed effects control for unobserved characteristics that are common to all observations within a dimension, e.g. country. This is an important feature since the data are likely to be plagued by unobserved heterogeneity across certain dimensions. Inclusion of fixed effects allows for different intercepts across countries for example, thus making it more likely that the model is correctly specified (Angrist & Pischke, 2009).

C. *ESS survey administration*

The European Social Survey is primarily focused on monitoring changing values and attitudes across Europe. The survey involves strict random probability sampling, a minimum target response rate of 70% and rigorous translation protocols. It is conducted by an hour-long face-to-face interview which includes questions on a variety of core topics and the survey also provides an extensive set of socio-structural "background" variables. (European Social Survey, 2014)

I will be using the sixth round, ESS6, which was conducted in 2012. The complete dataset is available at their website (European Social Survey, 2017). ESS6 is very suitable as it is one of the largest surveys containing answers from 24 EU countries and since it was conducted by the end of the 2007-2013 funding period, the potential effects of the funding should have been captured in the survey responses.

Complementing the ESS6 with other survey rounds would bring additional value to this study, however, due to the difference in the questionnaires and regional composition combined with very limited available data on regional funding this was not an option at the time of writing this.

One major limitation with using the European Social Survey in this study is the inconsistency of the regional level which the survey was conducted. In the vast majority of the countries the survey was conducted at NUTS2 region level but there are two noticeable exceptions, Germany and the United Kingdom, where it was conducted at NUTS1 level. This forces us to exclude these responses as they do not match with the level the funding was allocated and this significantly decreases the number of regions in the sample.

D. *Data description*

After matching the survey responses of the remaining regions with the regional economic data we are left with 169 regions in the full dataset. Each region typically has above 100 respondents and we can see a very large difference in allocated funds for those eligible for convergence funding, the treated, and those who receive funding under other objectives. The average support for EU unification is around 5 on the 0-10 scale for both treated and non-treated regions, although slightly higher for the treated.

Table 1. Data description

	Full dataset		Data +/- 25 points from the 75 % threshold	
	Treated regions	Non-treated regions	Treated regions	Non-treated regions
Number of NUTS2 regions	70	99	29	18
Total allocated funds, average (millions), per region	2088.48	144.65	1823.84	241.21
Average number of survey respondents per region	256	187	146	164
Total number of survey respondents	17,961	18,554	4238	2964
Average support for EU unification (scale 0-10)	5,26	4,81	5,27	4,90

3. Assessment of Internal Validity of Research Design

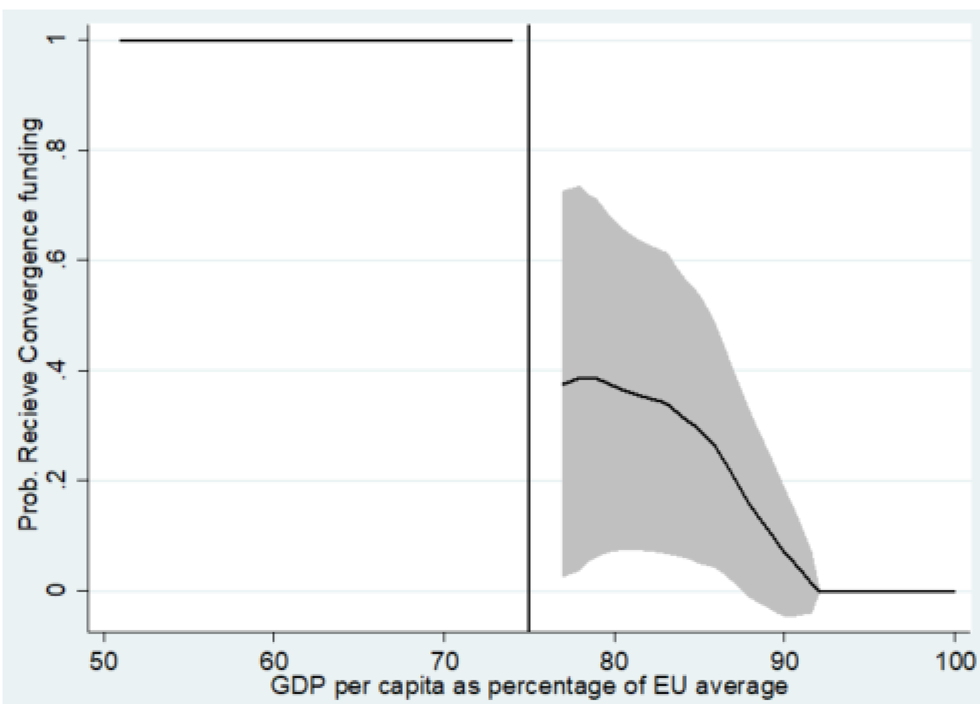
A. Discontinuities: First Stage and Reduced form

Imbens & Lemieux (2008) tell us that graphical analysis should be integral part of any RD analysis. Graphing the discontinuities and inspecting whether there is a clear jump or fall in the conditional mean around the threshold is a simple yet powerful way to visualize the identification strategy.

First stage

In order to implement the regression discontinuity approach, assignment to treatment must vary discontinuously at the threshold. Graph 1 presents a nonparametric plot of a region's probability of receiving convergence funding as a function of its GDP per capita as a percentage of the EU average, focusing on the narrow range of +/- 25 percentage points from the 75 percent threshold. The probability of receiving convergence funding is on the left vertical axis, and regional GDP per capital as a percentage of EU average is on the horizontal axis. The graph shows a clear drop in probability after the threshold.

Graph 1. Discontinuity in probability of treatment around the cutoff



All regions under threshold have a hundred percent probability of receiving funding. Regions above the threshold are significantly less likely to receive this fund although some do under the phasing-out condition which explains the nonzero probability. Just after the threshold, regions are 60 percentage points less likely to receive the funding and the difference increases further from the threshold.

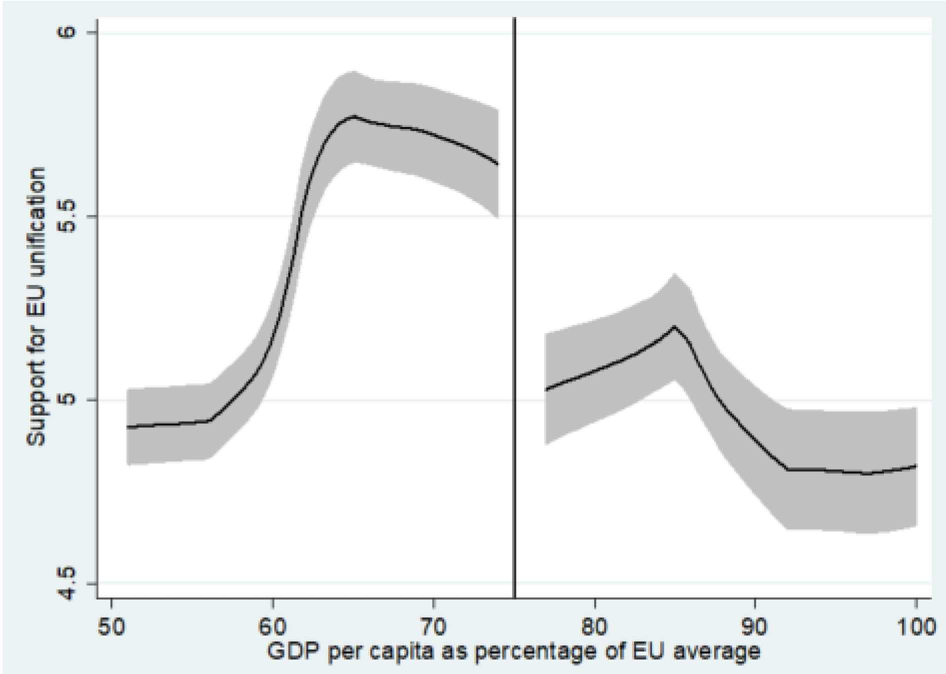
The non-perfect compliance confirms the need for a fuzzy design. Graph 1 essentially illustrates the first stage in an instrumental variable framework and suggests a strong first stage meaning that eligibility for funding is a valid instrument for funding itself. The complete estimation is done in Section 4.

In addition to the first stage defined as Equation (1) and presented in Graph 1, I also explored an alternative first stage that uses the discontinuity in the total amount of funds received by regions around the 75% threshold. We know that regions eligible for convergence funding on average receive a far larger sum of funds than those not eligible, hence we expect to see a clear discontinuity around the threshold. This alternative first stage, defined as Equation (A1), along with a graphical representation and estimated discontinuity is provided in Appendix A1. Despite that the graph shows a clear discontinuity in sum of allocated funds before and after the threshold, and regions eligible receive a far larger sum of funds than those not eligible, it still turned out to be an insignificant first stage when estimated together with the forcing variable, and will therefore not be used further in this thesis.

Reduced form

Graph 2 illustrates difference in support for EU unification around the threshold, i.e. the reduced form. Despite the somewhat odd shape of the curve it appears to be a clear fall after the threshold suggesting that regions that receive large amounts of funding are more supportive of further EU unification if they are not too far below the threshold. Again, the complete estimation is done in Section 4.

Graph 2. Difference in Support for EU unification around the cutoff



B. Continuity and Manipulation Checks

Like most other identification strategies, the regression discontinuity design requires that the treatment and control groups are similar with respect to their characteristics except for the treatment itself. If the funding is related to unobservable factors related to support for EU integration and if these factors cannot be controlled for, they will be absorbed by the error term and cause an omitted variable bias which subsequently means that the regression will not yield the causal effect. When a RD design is employed, the independence assumption is typically imposed on a narrow band around the treatment threshold (Angrist & Pischke, 2009). This means that regions just around the threshold should be very similar and comparable when it comes to all factors except the treatment. Table 2 presents the discontinuity on background characteristics of the survey respondents over a +/- 25 points bandwidth around the threshold. For the independence assumption to hold and the RD be valid, the discontinuity of the background characteristics should be small and insignificant, i.e. we need the background characteristics to be continuous over the threshold.

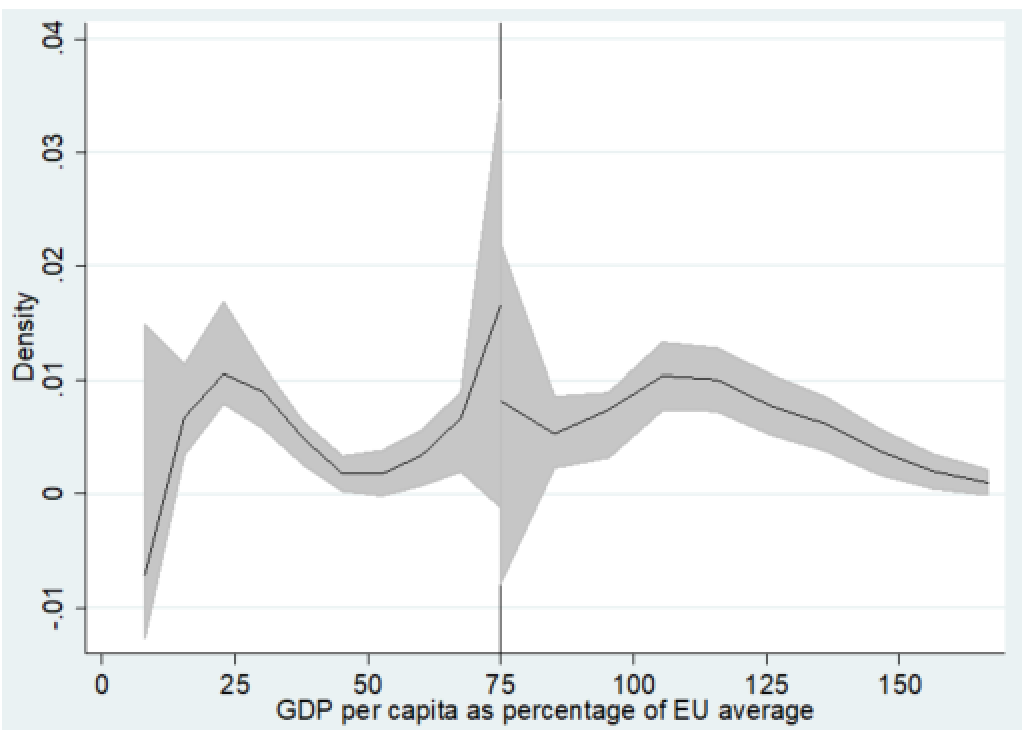
Table 2. Continuity of Background Characteristics

Satisfied with state of economy	.2467053 .1931739	Political interest	-.1784479 (.104131)	Ever had child in household	-.046779 .0966949
Satisfied with national gov.	-.270251 .2212314	Voted last national election	.00351799 .0601539	Household net income	.4248975 .4877547
Satisfied with state of democracy	-.1252756 .2810242	Left/Right political scale	-.2298862 .1837723	Years of education	1.300111* .6760626
Trust national parliament	.2668208 .1877879	Immigration bad/good for economy	.4238033 .2647146	Respondent happiness	-.5413367* (.2844885)
Trust in legal system	-.3176746 .2389926	Allow more immigration of different race	-.1988542** .0957585	Low skilled job	.322787 2.086977
Trust national political parties	-.1665609 .1644946	Immigration makes country worse/better	.5312257** .2473094	Have ever worked abroad	-.0072495 .0172902
Trust European Parliament	-.1866826 .1764227	Immigration bad/good for culture	.489671* .2723582	Have ever been unemployed more than 12 months	.1947258 .2589775
Traditions important	.0529048 .1798011	Gender	.007103 .0294257	Age	.6824523 2.536988

As is evident from Table 2, the majority of the background characteristics does not differ significantly over the threshold. However, some do, particularly those related to views on immigration. Failure to control for these covariates may threaten the internal validity of the research design or at the very least induce a bias in the results. Adding these covariates to the outcome equation can improve the precision of the estimate by eliminating bias, especially when using data further away from the threshold (Imbens & Lemieux, 2008). An important note on the views on immigration is that they actually might be an outcome of the convergence funding, hence should not be treated as background characteristics. Since three out of four variables related to immigration differs significantly over the threshold this is quite likely. An explanation for this is that the if funding has positive economic effects, as previous research suggests, then it is reasonable to expect that recipient regions also experience an increase in immigration or decrease in migration as they now are more economically attractive. This can in turn have an impact on the local population’s view on immigration. Therefore, when I in section 4 present results including background characteristics I do not include variables related to immigration as controls. The other characteristics will be included and we will later see that controlling for those does not make a sizeable impact on the results.

Another important check is the density of the assignment variable around the threshold. Lee & Lemieux (2010) and McCrary (2008) tell us that regression discontinuity designs can be invalid if the assignment variable, in this case GDP per capita of EU average, can be precisely manipulated by for example regional bureaucrats or statisticians working for the EU. A large jump or fall in density at the threshold likely implies sorting which threatens the RD design. When there is a big economic benefit of receiving these funds, one must be aware that this creates incentives to obtain these benefits, potentially by manipulating regional economic statistics. McCrary (2008) proposes a test based on an estimator for the discontinuity at the cutoff in the density function of the assignment variable. Cattaneo (2016) provides a manipulation tests following McCrary based on density discontinuity, constructed using the results for local polynomial distribution estimators.

Graph 3. Check for manipulation of assignment variable



Notes: Graph constructed using STATA rddensity plot command as provided by Cattaneo (2016)

Graph 3 presents the result of this test using the computed bandwidth of 30 on both sides of the threshold. The graph does show a slightly higher density just below the threshold, however, the estimated p-value, see appendix Table A2, of 0.4195 makes us not reject the null hypothesis of no manipulation thus the RD design is still valid in this regard. Graph A2 in the appendix provides a complete overview of the regional density in respect to GDP per capita as percentage of EU average. This too shows a slightly higher concentration of regions just before the threshold, but again, it is not enough for us to conclude that manipulation took place according to the manipulation test.

4. Estimation Results

A. *Baseline results*

The estimation is done taking a nonparametric approach, using only data close to the cutoff. When choosing bandwidth, one should be aware that there is a trade-off between bias and precision of the estimate. This is due to the fact that sample size is usually limited and thus makes it impossible to compare only individuals marginally close to the cutoff threshold, since this approach would yield a number of observations that is insufficient to draw precise inference. Choosing a larger bandwidth is necessary to find significant treatment effects, but tends to make the estimated effects less credible, since the assumption of similarity between treatment and control groups is more likely to be violated the further we move away from the cutoff. Using a nonparametric approach reduces the number of observations but alleviates functional form issues because a smaller bandwidth decreases the potential for specification bias and thus does not require an as detailed polynomial structure.

Table 3 contains the baseline first stage estimations and the reduced form estimations, and Table 4 the causal structural estimation. All regressions are for comparison run using both a local linear estimation and a local quadratic polynomial estimation. The results are also presented for two different bandwidths, +/- 25 and +/- 20 points away from the 75% threshold in GDP per capita as a percentage of EU average that determines whether a region is eligible or not for funding under the convergence objective.

Table 3. First stage and Reduced form estimates, Baseline

Linear functional form	+/- 25 points from the 75% threshold			+/- 20 points from the 75% threshold		
	(1) First stage, Treatm	(2) Reduced form, Unif.Supp.	(3) Reduced form, Unif.Supp.	(4) First stage, Treatm	(5) Reduced form, Unif.Supp.	(6) Reduced form, Unif.Supp.
Eligible	0.608***	0.874***	0.817***	0.544**	0.594*	0.521*
	(0.175)	(0.306)	(0.296)	(0.210)	(0.310)	(0.305)
GDPofavg	-0.00968*	0.0223	0.0203	-0.0139*	-0.00530	-0.00825
	(0.00485)	(0.0147)	(0.0138)	(0.00777)	(0.0152)	(0.0140)
Constant	1.0265*	3.0449**	2.799**	1.410*	5.385***	5.216***
	(0.489)	(1.204)	(1.169)	(0.736)	(1.267)	(1.241)
Observations	47	7,170	7,170	38	4,977	4,977
R^2	0.725	0.0074	0.0149	0.725	0.0147	0.0230
Prob>F	<0.001	0.0153	<0.001	<0.001	0.0084	<0.001
Controls	-	No	Yes	-	No	Yes
Fixed effects	No	No	No	No	No	No

Quadratic functional form	+/- 25 points from the 75% threshold			+/- 20 points from the 75% threshold		
	(7) First stage, Treatm	(8) Reduced form, Unif.Supp.	(8) Reduced form, Unif.Supp.	(10) First stage, Treatm	(11) Reduced form, Unif.Supp.	(12) Reduced form, Unif.Supp.
Eligible	0.532**	0.779**	0.725**	0.522**	0.682*	0.601*
	(0.206)	(0.339)	(0.329)	(0.221)	(0.363)	(0.355)
GDPofavg	0.0653**	0.190**	0.181**	0.0966**	0.175**	0.156*
	(0.0286)	(0.0845)	(0.0810)	(0.0431)	(0.0784)	(0.0800)
$GDPofavg^2$	-0.00051**	-0.00117**	-0.00113**	-0.00072**	-0.00114**	-0.00104*
	(0.000228)	(0.000544)	(0.000525)	(0.000328)	(0.000516)	(0.000525)
Constant	-1.565**	-2.610	-2.609	-2.713**	-1.628	-1.165
	(0.6844)	(3.321)	(3.147)	(1.204)	(3.00478)	(3.0283)
Observations	47	7,170	7,170	38	4,977	4,977
R^2	0.757	0.0142	0.0211	0.699	0.0178	0.0255
Prob>F	<0.001	0.0184	<0.001	<0.001	0.0032	<0.001
Controls	-	No	Yes	-	No	Yes
Fixed effects	No	No	No	No	No	No

Notes: Standard errors clustered at NUTS2-regional level in parentheses *** p<0.01, ** p<0.05, * p<0.1. No controls for respondent background characteristics in first stage estimations as they are estimated using regional level data.

In order to calculate the coefficient of the structural equation using 2SLS estimation I first apply Equation 4.

$$Structural\ form\ coefficient_{2SLS} = \frac{Reduced\ form\ coefficient}{First\ stage\ coefficient} \quad (4)$$

However, even if this is the proper way to calculate the IV-coefficient a problem arises due to the fact that the first stage is estimated using regional level data and the reduced form estimation uses individual level data. To find the standard errors and significance levels, we need to use the same level of aggregation. Table 4 presents the 2SLS results of using individual level aggregation and as we see the coefficient is fairly similar and significant at a conventional level. Since the aggregation level of the first stage is different in this estimation the yielded coefficients are not exactly those of Equation 4, they are however relatively similar and this was a necessary step in order to estimate the standard errors and significance level. For the reader I suggest to interpret the coefficients calculated using Equation 4 as the actual coefficients and to look at the treatment coefficient in Table 4 to get a relatively accurate idea of the standard errors and significance level. The effect of the funding ranges from about 0,8-1,5 units on the 0-10 scale in additional unification support, i.e. 8-15 percentage points.

Table 4. Structural form estimates, Baseline

	+/- 25 points from the 75% threshold			+/- 20 points from the 75% threshold		
	(1) 2SLS Linear functional form, Unif.Supp.	(2) 2SLS Quadratic functional form, Unif.Supp.	(2) 2SLS Quadratic functional form, Unif.Supp.	(3) 2SLS Linear functional form, Unif.Supp.	(2) 2SLS Quadratic functional form, Unif.Supp.	(4) 2SLS Quadratic functional form, Unif.Supp.
$\frac{Reduced\ form\ coefficient}{First\ stage\ coefficient}$	1.344	1.464	1.362	0.957	1.306	1.151
Treatment	1.139**	1.0485***	1.0451***	0.801**	0.982**	0.968**
	(0.472)	(0.388)	(0.391)	(0.342)	(0.363)	(0.359)
Observations	7,170	7,170	7,170	4,977	4,977	4,977
R^2	0.0100	0.0113	0.0147	0.0201	0.0163	0.0208
Prob>chi2	<0.001	0.0064	<0.001	<0.001	0.0003	<0.001
Controls	Yes	No	Yes	Yes	No	Yes
Fixed effects	No	No	No	No	No	No

B. *Robustness*

This section is dedicated to test the robustness of the baseline results, this is done by the inclusion of fixed effects and changing the bandwidth size around the threshold.

As previously mentioned, the inclusion of fixed effects is important if the data suffers from unobserved heterogeneity across certain dimensions as it allows for different intercepts thus making it more likely that the model specification is correct. When deciding at which level to use fixed effects one should try to identify the smallest dimension in which observations share common unobservable characteristics. In this data, this dimension would be NUTS2 regions as respondents within a region probably have some unobserved characteristics common to that particular region, however, NUTS2 fixed effects are not possible with this data as we only observe each region once meaning that the different intercepts would capture the entire variation. The next dimension would be country level but this also has some issues. In many countries, all regions are either above or below the convergence funding threshold. For example, in Bulgaria and Hungary all regions receive convergence funding, and in Sweden and Denmark non do. This means that country fixed effects might also capture too much of the variation. Splitting the countries up in Northern-, Eastern-, Southern-, and Western-Europe (N,E,S,W-Europe) would work better as it would capture unobserved characteristics in these different areas without capturing too much of the variation. The countries within these areas share many unobserved characteristics related to history, economy and politics, and N,E,S,W-Europe fixed effects would help control for these.

Table 5 contains the re-estimation of the first stage and reduced form while including columns for fixed both country- and N,E,S,W-Europe fixed effects.

Table 5. First stage and Reduced form estimates, Robustness to Fixed Effects

Linear functional form	+/- 25 points from the 75% threshold				+/- 20 points from the 75% threshold			
	(1) First stage, Treatm.	(2) First stage, Treatm.	(3) Reduced form, Unif.Supp.	(4) Reduced form, Unif.Supp.	(5) First stage Treatm.	(6) First stage, Treatm.	(7) Reduced form, Unif.Supp.	(8) Reduced form, Unif.Supp.
Eligible	0.538**	0.634***	0.0637	0.866***	0.473**	0.543**	-0.354	0.656**
	(0.231)	(0.181)	(0.256)	(0.307)	(0.252)	(0.211)	(0.354)	(0.333)
GDPofavg	-0.0130*	-0.00721	0.00357	0.0246	-0.0153*	-0.0137*	-0.0257**	-0.00239
	(0.00733)	(0.00585)	(0.00842)	(0.0180)	(0.00891)	(0.00779)	(0.0105)	(0.0171)
Constant	1.196*	0.947*	2.879***	1.281	1.409*	1.793**	5.527***	5.077***
	(0.674)	(0.514)	(0.760)	(1.732)	(0.820)	(0.751)	(0.955)	(1.614)
Observations	47	47	7,170	7,170	38	38	4,977	4,977
R^2	0.825	0.733	0.0430	0.0142	0.804	0.711	0.0445	0.0231
Prob>F	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
Controls	-	-	Yes	Yes	-	-	Yes	Yes
Fixed effects	Country	N,E,S,W-Europe	Country	N,E,S,W-Europe	Country	N,E,S,W-Europe	Country	N,E,S,W-Europe

Quadratic functional form	+/- 25 points from the 75% threshold				+/- 20 points from the 75% threshold			
	(9) First stage, Treatm	(10) First stage, Treatm	(11) Reduced form, Unif.Supp.	(12) Reduced form, Unif.Supp.	(13) First stage, Treatm	(14) First stage, Treatm	(15) Reduced form, Unif.Supp.	(16) Reduced form, Unif.Supp.
Eligible	0.465**	0.562***	-0.0913	0.768**	0.464**	0.519**	-0.364	0.797**
	(0.245)	(0.206)	(0.332)	(0.344)	(0.251)	(0.223)	(0.353)	(0.379)
GDPofavg	0.0689*	0.0812**	0.0611	0.197**	0.0821	0.0804*	-0.113	0.199**
	(0.0353)	(0.0330)	(0.0562)	(0.089)	(0.0353)	(0.045)	(0.0702)	(0.0899)
$GDPofavg^2$	-0.0005**	-0.0006**	-0.000421	-0.0012**	-0.00064	-0.00062*	0.000574	-0.0013**
	(0.00027)	(0.00024)	(0.00043)	(0.00056)	(0.00045)	(0.00035)	(0.00044)	(0.00057)
Constant	-1.677	-2.0737**	1.140	-2.782	-2.148	-1.721	8.702***	-2.956
	(0.991)	(0.902)	(1.588)	(3.449)	(2.208)	(1.392)	(2.748)	(3.719)
Observations	47	47	7,170	7,170	38	38	4,977	4,977
R^2	0.851	0.772	0.0432	0.0207	0.817	0.729	0.0446	0.0265
Prob>F	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
Controls	-	-	Yes	Yes	-	-	Yes	Yes
Fixed effects	Country	N,E,S,W-Europe	Country	N,E,S,W-Europe	Country	N,E,S,W-Europe	Country	N,E,S,W-Europe

Notes: Standard errors clustered at NUTS2-regional level in parentheses *** p<0.01, ** p<0.05, * p<0.1. No controls for respondent background characteristics in first stage estimations as they are estimated using regional level data.

When comparing the results from Table 5 to the baseline estimations in Table 3, we observe that the first stage remains significant using both country- and N,E,S,W-Europe fixed effects with a small drop in the size of the coefficient for country fixed effects and a very small increase under N,E,S,W-Europe fixed effects, which in turn lets us conclude that the first stage is robust to fixed effects.

Things are a little bit different with the reduced form. When including country fixed effects all significant effects disappear and the coefficient becomes completely unreliable even going between positive and negative. This is most likely due to the problems explained above, therefore we will shift attention to the N,E,S,W-Europe fixed effects which gives results very similar to those of the baseline estimates, suggesting that the reduced form is robust to these fixed effects. The choice of functional form and bandwidth does have an impact on the estimates but is quite small suggesting that the model is not very sensitive.

Table 6 presents coefficients calculated using Equation 4 as well as the 2SLS estimation using individual level data carried out to obtain the standard errors and significance level. As with the baseline, the difference between the coefficients are relatively small.

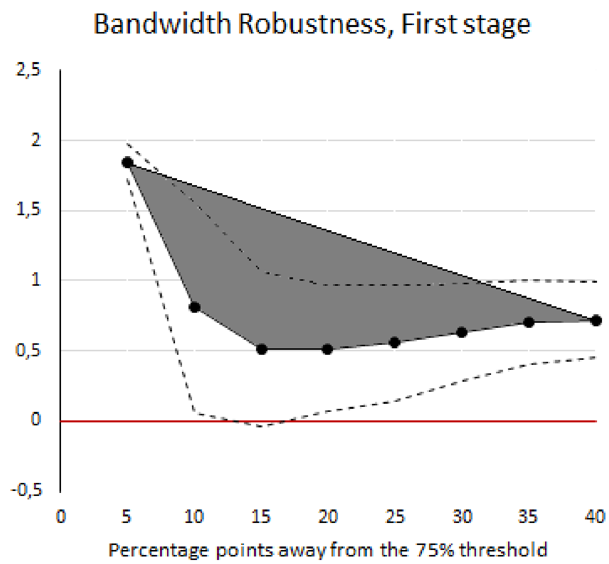
Table 6.. Structural form estimates, Robustness to Fixed Effects

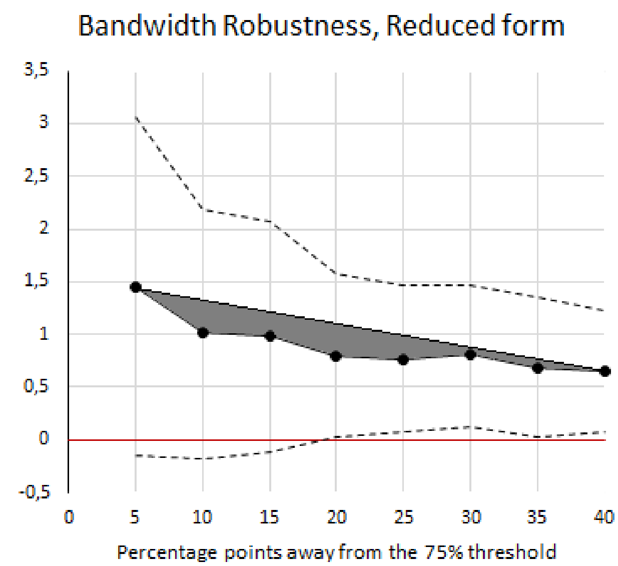
	+/- 25 points from the 75% threshold		+/- 20 points from the 75% threshold	
	(1) 2SLS Linear functional form, Unif.Supp.	(2) 2SLS Quadratic functional form, Unif.Supp.	(3) 2SLS Linear functional form, Unif.Supp.	(4) 2SLS Quadratic functional form, Unif.Supp.
$\frac{Reduced\ form\ coefficient}{First\ stage\ coefficient}$	1.365	1.366	1.208	1.535
Treatment	1.161**	1.158***	1.121**	1.311**
	(0.467)	(0.402)	(0.435)	(0.468)
Observations	7,170	7,170	4,977	4,977
R^2	0.0124	0.0172	0.0248	0.0261
Prob>chi2	<0.001	<0.001	<0.001	<0.001
Controls	Yes	Yes	Yes	Yes
Fixed effects	N,E,S,W-Europe	N,E,S,W-Europe	N,E,S,W-Europe	N,E,S,W-Europe

Further, to check how sensitive the results are to changes in bandwidth size, the first stage and reduced form were estimated using an additional 6 bandwidth sizes around the threshold.

Graph 4 and 5 illustrate how the size of the coefficient and confidence interval differ between the different bandwidths. Overall, the differences are small with the only exception being the +/-5 bandwidth for the first stage which is purely due to a statistical anomaly from the very small number of regions in that bandwidth and that those above the threshold are phasing-out regions.

Graph 4,5 First stage and Reduced form sensitivity to bandwidth size





Notes: Estimated using a quadratic functional form, N,E,S,W-Europe fixed effects and NUTS2 clustered standard errors. Controls were included in the reduced form. The solid black line represents the coefficient of interest. The area between the dotted lines represents the 95% confidence interval.

C. Spillover effects

When doing studies with regional treatment one must be aware of potential spillover effects. It could be that the effects of receiving convergence funding in one region spills over onto a neighboring region, in other words, the treatment effect might spill over into the control group. This could be a potential issue as it would lead to biased coefficients. This could be tested by creating a variable for proximity to treated region or a dummy for treatment neighbor as Becker & Egger (2010) did in their RD study of the effects on economic growth of the funding. Their estimates did not change to any larger extent as a result of controlling for spillovers and I do not think it would in this thesis either. The reason I believe spillover effects will not threaten the findings of this thesis is that even if there are any spillovers they would most likely increase the support of EU unification in non-treated regions as a result trade, opinions spreading over boarders etc. Therefore, as a result, we would expect the difference between non-treated and treated to become narrower. This, in turn, would imply that the true effect of the treatment would be larger than the estimated coefficient, thus rendering the estimated coefficient a lower bound of the true effect. Although a negative effect is possible if for example the control groups feel jealousy towards there fund receiving neighbors and this decreases their EU support, I find this scenario less likely.

5. Conclusion

This thesis used a fuzzy regression discontinuity design and found a positive effect of the EU Cohesion Policy funds on public support for EU integration and unification, leading us to not reject the hypothesis. Regions receiving funds under the convergence objective are on average 8-15 percentage points more in favor of EU integration and unification than the control regions, depending on the model specification and bandwidth. The result is in line with literature and is sufficiently robust to the inclusion of N,E,S,W-Europe fixed effects and changes in bandwidth size. The result of this thesis is comparable to the result of Massetti & Schakel (2016), although the effect I find on public support is considerably smaller than what they found on party positions.

Nevertheless, the result of this thesis adds to the importance of the Cohesion Policy as an instrument to keep the union together and further European integration. One should however be careful when drawing conclusions from this thesis. The results suggest a small to moderate positive effect but we should not lay too much weight on the exact coefficients as they do change slightly between the different specifications, bandwidths and estimations. Further, the small number of regions around the threshold is a considerable limitation. As more ESS rounds become available and the European Commission publish more of their Cohesion Policy data further research will have the potential to become more accurate.

References

Angrist, J. D., & Pischke, J.-S. (2009). *Mostly harmless econometrics : an empiricist's companion.* Princeton: Princeton University Press.

Armingeon, K., & Ceka, B. (2013). The loss of trust in the European Union during the great recession since 2007: The role of heuristics from the national political system. *European Union Politics, 15*(1), 82–107.

Becker, S., & Egger, P. (2010). Going NUTS: The effect of EU Structural Funds on regional performance. *Journal of Public Economics, 94*(9), 578-590.

Besley, T., & Coate, S. (2003). Centralized versus decentralized provision of local public goods: a political economy approach. *Journal of Public Economics, 87*, 2611-2637.

Bolton, P., & Roland, G. (1997). The Breakup of Nations: A Political Economy Analysis. *The Quarterly Journal of Economics, 112*(4), 1057-1090.

Cattaneo, M. D. (2016). rddensity: Manipulation testing based on density discontinuity. *The Stata Journal*(ii), 1-18.

Doyle, O., & Fidrmuc, J. (2006). Who favors enlargement?: Determinants of support for EU membership in the candidate countries' referenda. *European Journal of Political Economy, 22*(2), 520-543.

European Commission. (2008). *Working for the regions - EU Regional Policy 2007-2013.* Brussels: European Commission.

European Commission. (2014, June). *Europa.eu.* Retrieved from An introduction to EU Cohesion Policy: http://ec.europa.eu/regional_policy/sources/docgener/informat/basic/basic_2014_en.pdf

European Social Survey. (2014). *ESS Round 6 (2012/2013) Technical Report.* London: Centre for Comparative Social Surveys, City University London.

European Social Survey. (2017, 05 11). *European Social Survey*. Retrieved from Data Download: http://www.europeansocialsurvey.org/data/download.html?r=6

Eurostat. (2010). *European Regional and Urban Statistics Reference Guide.* Luxembourg: European Commision.

Eurostat. (2017, 05 11). *Eurostat*. Retrieved from Database: http://ec.europa.eu/eurostat/web/regions/data/database

Gabel, M., & Whitten, G. D. (1997). Economic Conditions, Economic Perceptions, and Public Support for European Integration. *Political Behaviour, 19*(1), 81-96.

Galgóczi, B., Leschke, J., & Watt, A. (2009). Intra-EU labour migration: flows, effects and policy responses. In B. Galgóczi, J. Leschke, & A. Watt, *EU Labour Migration Since Enlargement: Trends, Impacts and Policies* (pp. 1-41). Brussels: European Trade Union Institute.

Hooghe, L., & Marks, G. (2005). Calculation, Community and Cues. Public Opinion on European Integration. *European Union Politics, 6*(4), 419–443.

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

Inforegio. (2017, 05 11). *European Commission Regional Policy Inforegio*. Retrieved from Data for Research: http://ec.europa.eu/regional_policy/en/policy/evaluations/data-for-research/

Lee, D., & Lemieux, T. (2010). Regression Discontinuity Designs in Economics. *Journal of Economic Literature, 48*(2), 281–355.

Lefkofridi, Z., & Schmitter, P. C. (2015). Transcending or Descending? European Integration in Times of Crisis. *European Political Science Review, 7*(1), 3–22.

Lockwood, B. (2002). Distributive Politics and the Costs of Centralization. *The Review of Economic Studies, 69*(2), 313-337.

Lubbers, M., & Scheepers, P. (2007). Explanation of Political Euro-Scepticism at the Individual, Regional and National Levels. *European Societies, 9*(4), 643-669.

Massetti, E., & Schakel, A. (2016). Buying Support for Europe: The Impact of Cohesion Funds on Regionalist Parties' Positions on European Integration. In S. Piattoni, & L. Polverari, *Handbook on cohesion policy in the EU* (pp. 217-227). Cheltenham: Edward Elgar Publishing.

McCrary, J. (2008). Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of Econometrics, 142*(2), 698–714.

Nicoli, F. (2016). Hard-line Euroscepticism and the Eurocrisis: Evidence from a Panel Study of 108 Elections Across Europe. *Journal of Common Market Studies, 55*(2), 312-331.

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. *Regional Science, 92*(1), 217-233.

Pieńkowski, J., & Berkowitz, P. (2015). *Econometric assessments of Cohesion Policy growth effects: How to make them more relevant for policy makers?* European Commision.

Ramajo, J., Márquez, M. A., Hewings, G. J., & Salinas, M. M. (2008). Spatial heterogeneity and interregional spillovers in the European Union: Do cohesion policies encourage convergence across regions? *European Economic Review, 52*(3), 551-567.

Treib, O. (2014). The voter says no, but nobody listens: causes and consequences of the Eurosceptic vote in the 2014 European elections. *Journal of European Public Policy, 21*(10), 1541-1554.

Turner, J., Brown, R., & Tajfel, H. (1979). Social comparison and group interest in ingroup favouritism. *European Journal of Social Psychology, 9*(2), 187-204.

Wincott, D., Peterson, J., & Convery, A. (2017). Introduction: Studying Brexit's causes and consequences. *The British Journal of Politics and International Relations, 19*(3), 429–433.

Appendix

A1. Alternative first stage

$$SumAllocatedFunds_{jk} = \beta_0 + \beta_1 ELIGIBLE_{jk} + f(GDPofavg_{jk}) + \varepsilon_{ijk} \quad \text{(A1)}$$

Graph A1. Alternative first stage, discontinuity in sum of funds, +/- 25 points from threshold

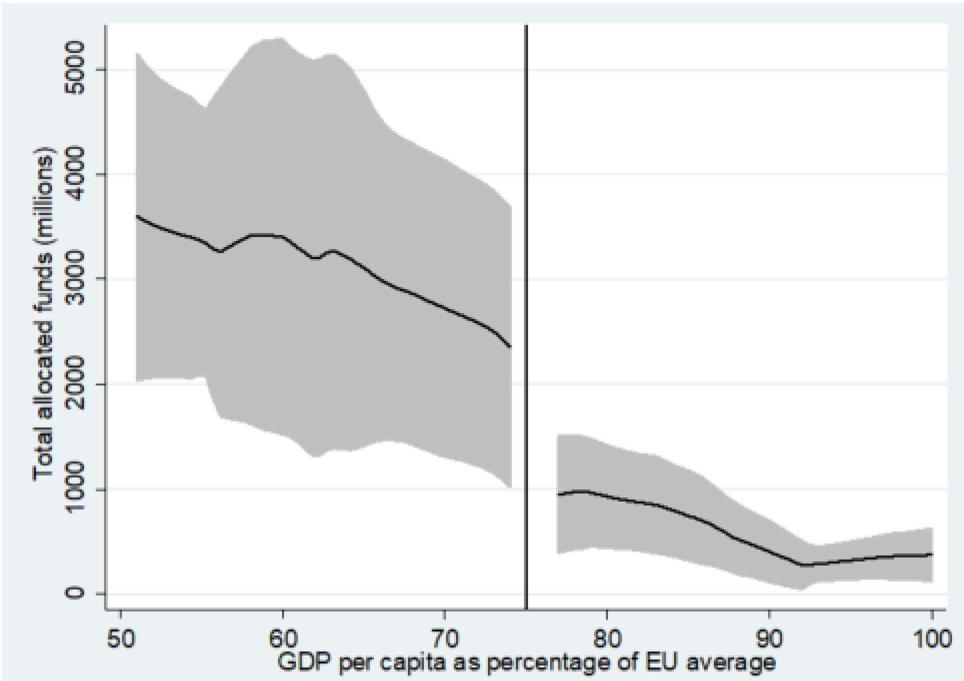


Table A1. Alternative first stage estimation, +/- 25 points from threshold

Linear regression

Number of obs	=	47
F(2, 46)	=	12.09
Prob > F	=	0.0001
R-squared	=	0.2894
Root MSE	=	1959

(Std. Err. adjusted for 47 clusters in region1)

sumallocat~s	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
eligible	1100.786	983.591	1.12	0.269	-879.0797	3080.652
GDPofavg	-55.1403	33.47711	-1.65	0.106	-122.5262	12.24564
_cons	5416.037	2940.365	1.84	0.072	-502.6116	11334.69

A2. Manipulation checks

Table A2. RD Manipulation Test

RD Manipulation Test using local polynomial density estimation.

Cutoff c = 75	Left of **c**	Right of **c**
Number of obs	**64**	**105**
Eff. Number of obs	**24**	**31**
Order est. (p)	**3**	**3**
Order bias (q)	**4**	**4**
BW est. (h)	**29.863**	**30.549**

Number of obs	=	**169**
Model	=	**unrestricted**
BW method	=	**comb**
Kernel	=	**triangular**
VCE method	=	**jackknife**

Running variable: GDPofavg.

Method	T	P>|T|
Robust	**-0.8072**	**0.4195**

Graph A2. Histogram, Regional density by GDP per capita as percentage of EU average