

ePub^{WU} Institutional Repository

Julia Bachtrögler and Christoph Hammer and Florian Schwendinger and Wolf Heinrich Reuter

Guide to the galaxy of EU regional funds recipients: evidence from new data

Article (Published)
(Refereed)

Original Citation:

Bachtrögler, Julia and Hammer, Christoph and Schwendinger, Florian and Reuter, Wolf Heinrich (2019) Guide to the galaxy of EU regional funds recipients: evidence from new data. *Empirica*, 46 (1). pp. 103-150. ISSN 0340-8744

This version is available at: <http://epub.wu.ac.at/6687/>

Available in ePub^{WU}: November 2018

ePub^{WU}, the institutional repository of the WU Vienna University of Economics and Business, is provided by the University Library and the IT-Services. The aim is to enable open access to the scholarly output of the WU.

This document is the publisher-created published version.



Guide to the galaxy of EU regional funds recipients: evidence from new data

Julia Bachtrögler^{1,2} · Christoph Hammer² · Wolf Heinrich Reuter³ · Florian Schwendinger²

© The Author(s) 2018

Abstract

This study presents a new firm- and project-level dataset containing data on over two million projects co-funded by the EU structural and cohesion funds in 25 EU member states during the programming period 2007–2013. Information on individual beneficiary firms and institutions is linked with business data of Bureau van Dijk's ORBIS database. Moreover, text mining techniques are applied to categorise the EU cohesion policy projects into fifteen thematic categories. Stylised facts reveal substantial regional heterogeneity in the distribution of funds to certain projects and beneficiaries (with respect to their size or industry). Furthermore, regional funds distribution differs across less developed and higher-income as well as urban and rural regions. In an econometric analysis, we control for project and firm characteristics that we expect to determine the single project's value, which is confirmed by the results. Nevertheless, there remains unexplained variation in individual project volumes, which differs systematically across countries.

Keywords Distribution of EU structural funds · Cohesion policy · Firm-level data · Cohesion · European Union

JEL Classification E61 · H77 · R11 · R58

1 Introduction

Under the cohesion policy framework, the European Union (EU) committed over EUR 348 billion from 2007 to 2013 (over EUR 371 billion in the multi-annual financial framework 2014–2020) to the re-distribution of funds among European regions

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s10663-018-9427-5>) contains supplementary material, which is available to authorized users.

✉ Julia Bachtrögler
julia.bachtroegler@wifo.ac.at

Extended author information available on the last page of the article

in order to foster regional development and cohesion. There is a broad literature that investigates different aspects of the effectiveness of those financial contributions, e.g., in terms of increasing income growth (see Dall’Erba and Fang 2017, for a survey). Due to a lack of data, research so far has mainly focused on the implementation of the policy and its effects at an aggregated, mostly regional, level. Nevertheless, the specific design of cohesion policy programmes remains a prominent theme in the academic and public debate, especially with respect to post-2020 EU cohesion policy. Apart from the allocation of funds to thematic priorities in a region, however, the intraregional distribution of funds to specific projects and beneficiaries has been a “black box” to researchers and European policy makers so far.

In principle, the European Regional Development Fund (ERDF), the European Social Fund (ESF) and the Cohesion Fund (CF), in this study together referred to as the EU’s *regional funds*, co-finance projects that are part of operational programmes (OP) and pursue strategic priorities like strengthening the labour market, improving social infrastructure or building better traffic networks.¹ The projects are carried out by firms, institutions or other entities and are selected to be co-financed by the particular OP’s managing authority (a public or private body nominated by the member state).² Therefore, next to the managing authorities’ responsibility for choosing suitable and promising projects, the actual beneficiaries are accountable for the single projects’ success. The appropriate fulfilment of these tasks (project selection and successful implementation) likely contributes to achieving the corresponding OP’s target and, finally, the overall effectiveness of a region’s EU cohesion policy implementation.³

This study presents newly collected information on individual projects co-financed by regional funds during the multi-annual financial framework (MFF) 2007–2013 and explores how the regional funds committed to European NUTS-2 regions are distributed within the regions.⁴ The resulting database contains data on over two millions of co-funded projects in 25 EU member states. Those projects are carried out by 1,076,097 beneficiaries which we matched with the ORBIS business database by Bureau van Dijk in order to gain information on their business characteristics.

Using that data, this paper contributes to the literature by providing a comprehensive analysis of the projects selected by national or regional managing authorities during the MFF 2007–2013. The high level of granularity of the data allows to investigate whether the distribution of regional funds differs across European

¹ We do not consider projects co-funded by the European Agricultural Fund for Rural Development (EAFRD) and the European Maritime and Fisheries Fund (EMFF) which are also EU cohesion policy instruments.

² Refer to European Council (2006b) for detailed information on cohesion policy implementation.

³ The European Commission considers various indicators, e.g., the number of jobs created or the number of direct investment aid projects to small and medium-sized enterprises, for its evaluation of cohesion policy (see http://ec.europa.eu/regional_policy/en/policy/evaluations/ec/2007-2013/#1).

⁴ Since the multi-annual financial framework 2007–2013, the managing authorities of the OPs are required to report the firms and institutions which receive financial support for carrying out projects that contribute to economic and social cohesion across European regions.

regions with respect to project, geographical and firm-level characteristics of the funds' beneficiaries.⁵ Thus, the availability of this micro data increases the transparency of policy implementation. For example, it shows whether the EU's funding priorities (e.g., the focus on small and medium-sized enterprises) are mirrored in cohesion policy implementation or if more guidance would be desirable. Moreover, the findings of this paper may be interesting for EU cohesion policy evaluators, as differences in the distribution of regional funds within regions may play a role for explaining differences in EU cohesion policy effectiveness across regions.

First, this study investigates the intraregional distribution of regional funds along different dimensions like the thematic priorities set by the OPs' managing authorities. This descriptive analysis distinguishes between different types of regions, i.e., less developed and richer, as well as urban and rural ones. One result is that a large share of regional funds in low-income regions is allocated to *Transportation Infrastructure, Environment and Innovation and Research and Technological Development (RTD)* projects. In the majority of the other regions, the largest project amounts are also assigned to the latter two themes, while, in addition, there is a stronger focus on labour market projects.

Second, since we are one of few contributions so far that are able to analyse regional fund data at the level of projects and beneficiaries in an encompassing way, we also focus on individual projects. In particular, we document substantial variation in project volumes across different funding instruments (i.e., the ERDF, ESF or CF), the objective of the OP (i.e., *Convergence* or *Regional competitiveness and employment*), the regional focus on thematic priorities, and beneficiaries' characteristics like firm size or industry. Moreover, related to the projects' themes, we find that managing authorities in urban (as compared to rural) and less developed (as compared to higher-income) regions allocate the regional funds on average to relatively large projects.

Third, after controlling for the project and beneficiary characteristics in a regression analysis, results show residual variation in project volumes that varies across countries and regions. In general, there is little literature on whether the size of (cohesion policy) grants to firms is related to the effectiveness and efficiency of a policy. Locatelli et al. (2017) indicate that the tendering of larger projects favours corruption in the public procurement process. Therefore, we argue that future research should investigate whether regional residual variation in average project values is linked to institutional settings or historically grown funding traditions in the respective country. Furthermore, the link between the latter as well as regional funds distribution patterns and the effectiveness of policy implementation should be explored.⁶

⁵ This paper provides details on the creation of the dataset that may be interesting to researchers working with similar data. One important feature in this context is the publication of the R package (*fastTextR*) which we use for classifying projects into fifteen themes defined by the European Commission and which we made publicly available.

⁶ E.g., for the field of industrial policy, Criscuolo et al. (2012) indicate that granting funds to smaller firms yields better results than funding large firms, as, beside other reasons, the latter are more likely to displace own investments by the firm grants. Bachtrögl et al. (2018) show for seven EU member states that, independently of their volume, projects carried out by manufacturing firms in regions with

The remainder of this paper is structured as follows: Chapter 2 gives an overview on the EU cohesion policy design and places our analysis into the context of existing literature. Chapter 3 describes the content of the database in detail and compares it with aggregate official data. Chapter 4 provides stylised facts regarding the distribution of total project volumes across (different types of) regions, project and firm characteristics. Chapter 5 shows the results of the econometric analysis with a focus on the individual project volume and, finally, Chapter 6 concludes.

2 The design of EU cohesion policy

In the multi-annual financial framework (MFF) 2007–2013, the regional funds amounting to EUR 348,865 million were the second largest item of the EU budget.⁷ The cohesion policy's objective is to increase convergence across European regions by co-financing member states' initiatives targeted at specific priorities. First, a NUTS-2 region's principle eligibility for funding is determined according to three main objectives, namely, (1) *Convergence* (former Objective 1), (2) *Regional Competitiveness and Employment* (former Objective 2), and (3) *Territorial Cooperation* (former Objective 3).⁸ Second, for eligible areas, the funds are allocated to operational programmes which co-finance projects that are carried out by private or public firms or organizations. The new dataset presented in this study is based on lists of these beneficiary firms, institutions, non-governmental organizations or other types of entities (in the following we refer to all types of beneficiaries by the term *firm*) that have to be made public since the MFF 2007–2013 (Article 7 in European Commission 2006).

2.1 The distribution of EU regional funds

The EU's cohesion policy works under the principle of shared management (refer to Article 14 and 15 in European Commission 2006). That is why multiple European and national (as well as sub-national) institutions are involved in the allocation process of regional funds.

Footnote 6 (continued)

lower GDP per capita tend to be (partly statistically significantly) more effective in increasing beneficiary firms' employment and value added growth than similar projects run by firms in richer regions within the same country.

⁷ Most expenditure, i.e., EUR 412,611 million, is allocated to the "natural resources" programme which includes agricultural subsidies. See http://ec.europa.eu/budget/figures/fin_fwk0713/fw0713_en.cfm#cf07_13.

⁸ NUTS regions at different levels are defined according to the *Nomenclature des unités territoriales statistiques* 2010 (NUTS 2010) (European Commission 2011). In the MFF 2007–2013, regions with a gross domestic product (GDP) per capita below 75% of the EU-25 average in 2000–2002 are eligible for funds under the *Convergence* objective. The remaining regions are eligible for transfers under the *Regional Competitiveness and Employment* objective and *Territorial cooperation* initiatives (Article 3 in European Council 2006b). Member states can apply for support from the Cohesion Fund if their gross national income (GNI) per capita lies below 90% of the EU average.

For each MFF, the European Council prepares a document with strategic guidelines on reducing economic, social and territorial disparities. These guidelines for 2007–2013 encompass three priorities (European Council 2006a): (1) improving transport infrastructure, environmental and energy issues, (2) creating more and better jobs, and (3) a focus on knowledge transfer and innovation. Regarding the latter, special emphasis is put on supporting small and medium-sized enterprises (SME) which “often represent the highest source of employment at the regional level” (European Council 2006a, p. 19). The strategic guidelines on cohesion serve as a basis for the so-called national strategic reference framework that needs to be provided by each member state. The national strategic reference framework provides an overview of fields for intervention of cohesion policy in the particular country and undergoes a review process by the European Commission. Besides a proposal for the annual allocation of regional funds across the period, the member states need to generate a list of operational programmes (per objective and fund) for the objectives *Convergence* and *Regional Competitiveness and Employment*. Each operational programme is prepared and implemented by a managing authority, a private or public body appointed by the member state (Article 59 of European Council 2006b). In the most cases (in the case of a national OP always) it has a specified thematic target, e.g., improving regional human capital or social infrastructure. In most cases, they refer to particular NUTS-2 or NUTS-1 regions, however, there are also national, NUTS-0, programmes (see Appendix A.1 and Title III in European Council 2006b). The OPs must incorporate reasons for focusing on specific priority axes proposed by the Commission (see Annex IV of European Council 2006b).⁹ From these priority axes, the European Commission derives fifteen so-called (priority) themes.

In the next step, the OPs’ managing authorities select appropriate projects that are carried out by firms, i.e., the beneficiaries.¹⁰ According to Article 2 of Council Regulation (EC) No. 1083/2006, a beneficiary is defined as “an operator, body or firm, whether public or private, responsible for initiating and implementing operations”. An operation is referred to as “a project or groups of projects selected by the managing authority of the operational programme [...] allowing achievement of the goals of the priority axis to which it relates” (European Council 2006b). The amount of EU co-funding for each project depends on the eligible expenditure and the designated co-financing rate.¹¹

The structure of the new database builds on these regulations. The dataset includes the OP to which each observation (project) is assigned, the corresponding fund, objective as well as the theme. Thus, we are able to check the validity of our

⁹ See also “ERDF/ESF/CF Priority theme overview 2007–2013” at http://ec.europa.eu/regional_policy/en/policy/evaluations/data-for-research/.

¹⁰ Note that data on applications for cohesion policy projects are not available. Therefore, we cannot control for the self-selection of firms.

¹¹ Annex III of European Council (2006b) reports the ceilings for co-financing rates, i.e., the maximum percentage of eligible expenditure that is financed by a regional fund. E.g., the maximum co-financing rate for Spain amounts to 80% for the *Convergence* and to 50% for the *Regional Competitiveness and Employment* objective. The detailed regulation on the eligibility of expenditure can be found in Council Regulation (EC) No. 1083/2006 and Commission Regulation (EC) No 1828/2006.

data by comparing it in different dimensions with official numbers on a more aggregate level by DG REGIO (see Sect. 3.3).

While the next section covers the literature that evaluates the effects of cohesion policy on regions or firms, there is also a literature that discusses the allocation process and the factors that potentially influence this bargaining. Bachtler and Mendez (2007) discuss the history of the allocation process up to the beginning of the MFF 2007–2013. They narratively highlight the negotiation process between member states and the EU Commission as well as the ongoing spatial concentration of the majority of funds. A more quantitative line of research (e.g., Bouvet and Dall’Erba 2010; Dellmuth 2011; Dellmuth and Stoffel 2012; Tosun 2014) finds that, indeed, factors such as the political situation, the type of governance, previous success of implementation or even the degree of euroscepticism drive part of the allocation process.

2.2 Literature review

Most studies that analyse cohesion policy in a pan-European setting focus on the regional (in many cases NUTS-2) level (Hagen and Mohl 2009; Pienkowski and Berkowitz 2016). Most recently, Dall’Erba and Fang (2017) provide a meta-analysis of econometric studies on the evaluation of cohesion policy effects.

Generally speaking, there is no consensus in the literature regarding the outcome of cohesion policy. The effectiveness is measured, e.g., as a positive effect on gross domestic product (GDP) per capita growth (e.g., Pellegrini et al. 2013), investments per capita (Becker et al. 2013) or regional research and development activity (Ferrara et al. 2016). Most studies find a conditional positive effect of regional funds assignment (e.g., Becker et al. 2013; Cappelen et al. 2003; Ferrara et al. 2016), while others provide results that even suggest a negative impact (Breidenbach et al. 2016). In recent years, the potential reasons for heterogeneous (conditional) cohesion policy effects have gained major attention.

First, Rodríguez-Pose and Fratesi (2004) move the focus to expenditure categories of the main funding instruments. They find that investments in infrastructure or agriculture do not have sustainable effects on regional growth (see also Puga 2002; Dall’Erba and Le Gallo 2008), though, projects that foster human capital lead to sustainable positive effects on economic cohesion. This and other studies use data on the distribution of expenditure across NUTS-2 regions and themes (e.g., Dall’Erba and Le Gallo 2007; Percoco 2013; Ferrara et al. 2016).

Second, regional heterogeneity as a determinant of policy effectiveness has become a topic of interest and is often modelled by a region’s capacity to take advantage of regional funds. Becker et al. (2013) indicate that human capital and institutional quality matter for the effectiveness of Objective 1 funds in terms of their effect on GDP per capita growth and investment. Institutions are confirmed as influencing factor for the success of cohesion policy by other authors as well (e.g., Cappelen et al. 2003; Bachtler et al. 2014; Rodríguez-Pose 2013). Recently, Gagliardi and Percoco (2017) show that European cohesion policy is most effective in rural regions that are located close to cities.

Third, Becker et al. (2012) take the amount of regional funds expenditure spent in a region into account (instead of treatment dummies) and conclude that there is a maximum efficient level of funds and paying more does not increase the effectiveness any more (see also Kyriacou and Roca-Sagalés 2012; Rodríguez-Pose and Garcilazo 2015).

Fourth, Becker et al. (2018) and Bachtrögler (2016) analyse the effects of structural funds (on income growth) in lagging regions over time and in the context of the economic and financial crisis starting in 2007. The latter finds that the effectiveness of cohesion policy in terms of increasing GDP per capita growth appears to decrease in the crisis compared to former periods when controlling for regional structural characteristics. Barone et al. (2016) also take the time dimension into account and show that cohesion policy effects are not persistent over time.

Finally, spatial heterogeneity and spillovers are considered to play a role for cohesion policy effectiveness (Le Gallo et al. 2011; Breidenbach et al. 2016; Maynou et al. 2016). This strand of the literature widely confirms some small positive effects on regional growth or convergence for a number of regions but no general overall effect.

The majority of studies named above are based on data at the regional or local level to study growth or convergence effects. One major drawback of this type of analysis is the potential endogeneity of structural funds. As they are especially granted to lower-income regions, cohesion policy is most likely not exogenous with respect to regional growth. Attempts to overcome this statistical problem include the use of time lags (Rodríguez-Pose and Fratesi 2004), different instruments (Dall'Erba and Le Gallo 2007, 2008), generalised method of moments (GMM) estimators (Breidenbach et al. 2016) or a computable general equilibrium approach (Horridge and Rokicki 2018). Another way to possibly identify causal effects is the use of microeconomic methods with regional or micro data.

Turning to the beneficiaries as unit of observation, De Zwaan and Merlevede (2013) evaluate the effects of Objective 1 and Objective 2 payments in the programming period 2000–2006 in 25 EU member states on productivity and employment growth of firms in treated and non-treated regions. However, they do not use actual recipients of regional funds but compare all manufacturing firms (available in the ORBIS database) located in treated regions with the manufacturing firms in non-treated regions. Additional firm-level analyses are available for sub-national geographical units (Bernini and Pellegrini 2011) and certain types of regional funds (Hartsenko and Sauga 2012).

As we overcome this lack of data with the new database, this paper contributes to the literature by giving the first detailed insights on actual beneficiaries (and projects) of cohesion policy in 25 EU member states between 2007 and 2013. A combined analysis of the projects' theme, firm-level characteristics of corresponding beneficiaries and the size of the projects' expenditure may help to explain heterogeneous effects of regional funds allocation found in the literature, and thereby, lead to important policy implications. Moreover, it allows to identify regional funds allocation patterns across regions, e.g., less developed and others as well as urban and rural regions, and countries (see Sects. 4 and 5).

3 A novel dataset

3.1 Content of the dataset

The European Commission's Directorate-General for Regional and Urban Policy (DG REGIO) provides a collection of links to national or regional websites that make lists of beneficiaries available.¹² Unfortunately, the degree of detail of these lists' content as well as their structure vary significantly across countries, regions and even operational programmes.¹³ Moreover, most documents are provided in national languages, different data formats and using non-standardised definitions. Besides collecting and processing all information, we extend the data on beneficiaries by matching it with the ORBIS business database by Bureau van Dijk.

The resulting set of variables can be grouped into three blocks, namely, (1) project information, (2) funding (co-financing) information, and (3) business characteristics of the beneficiary retrieved from ORBIS. First, the project information includes the country and NUTS region in which the project is carried out according to the OP and list of beneficiaries, respectively.¹⁴ Moreover, it covers the co-financing fund, the objective and corresponding OP to which the project is assigned. As already noted, the dataset includes projects co-funded by the ERDF, the ESF and the CF, under the objectives of *Convergence* and *Regional Competitiveness and Employment*. Next, a project name or description, the start and end date as well as the theme of the project are specified. The theme is not reported by all managing authorities, which is why we classify the remaining projects according to available project information using supervised text classification (see Sect. 3.2 for a detailed description). The fifteen themes are: (1) *Capacity Building*, (2) *Culture, Heritage and Tourism*, (3) *Energy*, (4) *Environment*, (5) *Human Capital*, (6) *Innovation & Research and Technological Development (RTD)*, (7) *IT Services and Infrastructure*, (8) *Labour Market*, (9) *Other SME and Business Support*, (10) *Other Transport*, (11) *Rail*, (12) *Road*, (13) *Social Inclusion*, (14) *Social Infrastructure*, and (15) *Urban and Territorial Dimension*.

The second group of variables describes the funding structure. It contains the committed co-financing amounts by the EU (C_EU) and the national public funding (C_NAT; including co-funding of the recipient regions) as defined in the beginning of the MFF 2007–2013.¹⁵ If the amount borne by the firm itself (ineligible cost, Inelig¹⁶) is reported, it is added to the sum of co-funding commitments in order to calculate a total value for project i :

¹² See http://ec.europa.eu/regional_policy/en/atlas/beneficiaries/.

¹³ In Appendix A.1, we provide an overview of all OPs together with information on each one covered in our database.

¹⁴ If that information is not evident in the managing authority's report, we consider the NUTS-2 region in which the beneficiary is located according to ORBIS, if available.

¹⁵ Appendix A.1 reports the degree of detail in which the financial information on projects is provided in the lists of beneficiaries.

¹⁶ Article 56 in European Council (2006b) states the definition of project expenditure that is eligible for co-funding.

$$\text{Total value}_i = C_EU_i + C_NAT_i + \text{Inelig}_i \quad (1)$$

In addition to the commitments, project values that were actually paid out by European (Paid_EU) or national (Paid_NAT) public funds are available for a subset of observations. If only the actually paid-out value is declared, the total project value represents the sum of EU (ERDF, ESF or CF) and national payments (including those from regional governments):

$$\text{Total value}_i = \text{Paid_EU}_i + \text{Paid_NAT}_i \quad (2)$$

Furthermore, the declaration date refers to the time of reporting of the respective list of beneficiaries. In case it is not noted, we use the date of download.¹⁷

The third information block relates to the beneficiary. This data is produced by a matching exercise (using the name of the firm and its home country) with the ORBIS business database. We are aware of several shortcomings of this database (see, e.g., Kalemli-Ozcan et al. 2015), however, it represents the most comprehensive and accessible international business database.¹⁸ The resulting dataset contains the firm's name in ORBIS and its location, its founding year and information on the industry in which it operates (NACE Rev. 2 industry and four-digits code), the firms' number of employees and sales volume. Moreover, there is a size classification by ORBIS that is based on at least one of the following variables: the firms' number of employees, total assets, operating revenue and whether it is listed at the stock exchange.¹⁹ Furthermore, the database contains information on whether a firm belongs to a corporate group and, if so, on the number of entities in this group.

Table 1 shows all variables and their coverage in the database, i.e., the share of all observations for which it is available. The OP, its location, the funding instrument, the objective as well as the name of the project and the beneficiary is available for each project. Moreover, the dataset provides at least a total project value for each observation. 39% of the observations could be matched with ORBIS.

For analysing whether the distribution of regional funds in (less) developed as well as in urban and rural regions, respectively, shows different patterns, we need to gather further regional characteristics. Less developed regions are defined as NUTS-2 regions which are eligible for funds under the *Convergence* objective (former Objective 1), i.e., whose income per capita is lower than 75% of the EU-25 average (in 2000–2002) (European Council 2006b). According to the share of the

¹⁷ Payments corresponding to the MFF 2007–2013 have been transferred until the end of 2015, that is, not all have been reported at the time of data collection.

¹⁸ More information on ORBIS is available on the following website: <http://www.bvdinfo.com/en-gb/our-products/company-information/international-products/orbis>.

¹⁹ Very large company: listed, operating revenue greater or equal to EUR 100 million or total assets greater or equal to EUR 200 million or 1000 or more employees. Large company: Operating revenue greater or equal to EUR 10 million or total assets greater or equal to EUR 20 million or 150 or more employees. Medium-sized company: operating revenue greater or equal to EUR 1 million or total assets greater or equal to EUR 2 million or 15 or more employees. Small company: companies that do not fall in any of the other categories. As companies are classified as small ones if no data points are available, we do not consider this variable in regressions in Sect. 5. However, comparing the distribution of the ORBIS size category with the number of employees (more than 15, 150 or 1000) shows a similar picture.

population living in urban areas, DG REGIO classifies European NUTS-3 regions into predominantly urban, predominantly rural and intermediate ones (see Dijkstra and Poelman 2011)]. In this study, we distinguish between predominantly urban and other regions. As one can see in Table 1, due to the low coverage of the more disaggregated NUTS-3 locational information in ORBIS, this variable is known for a third of all observations.

3.2 Missing themes

As described in Sect. 2, the projects can be categorised into fifteen themes as defined by DG REGIO. Since not all managing authorities publish these themes in their lists of beneficiaries but most of them provide a project description and a project name, we employ supervised text classification to predict the missing project themes.

In order to train the classification algorithm (classifier) we use theme labels reported by some managing authorities, augmented with manually assigned theme labels. Since some of the project descriptions are given in a language other than English, we first use Google Cloud Translation API to translate the project descriptions and project names into English. Although we cannot quantify how many errors have been introduced during the translation process, we report the overall accuracy of the classification, where some part of error is attributed to translation errors. The records which cannot be translated are left unchanged. Then we remove those observations where the project name and description together have fewer than 30 characters. Thus, we overall use 1,698,191 projects (82.62% of all observations) and 588,713 labeled projects (28.64% of all observations) to train and evaluate the classifier.

Choosing an appropriate classifier is a non-trivial task. We find the recently published `fastText`²⁰ library (Joulin et al. 2016) to perform well on our dataset in terms of the performance metrics precision, recall and accuracy.²¹ In text classification, it is often desirable not to use the words of a text directly for estimation but to first map the text into a vector space with a much lower dimension. The `fastText` library which uses a single hidden layer neural network can be applied for text classification and to learn the vector representations of words.

The basic idea of the model is to proceed in two steps: first, the data is mapped into a low dimensional vector space (i.e., each sentence is mapped into a numeric vector) in such a way that similar texts have similar vector representations. Second, multinomial logistic regression is used to predict the labels.

For the evaluation of the classification we use the tenfold cross validation method (Stone 1974). In k -fold cross validation the data is randomly split into k parts, where $k - 1$ parts are used for training the model and the remaining part is used for model evaluation. To test the model on all available data, this process is typically repeated k times. In Fig. 1 we report the confusion matrix (a special type of contingency

²⁰ <https://github.com/facebookresearch/fastText>.

²¹ Due to the size of our data some of the alternative methods are overly time-consuming or run out of memory and are therefore excluded from further considerations.

Table 1 List of variables and their coverage in the dataset

Variable	Coverage (%)
<i>Project information</i>	
Country (NUTS-0)	100
Location (NUTS-2, NUTS-1)	97
Type of fund (ERDF, ESF, CF)	99
Objective (unique)	75
Operational programme	99
Name or description of project	97
Detailed project description	39
Project start date	19
Project end date	14
Theme	83
Name of the beneficiary	100
<i>Funding information</i>	
Currency	100
Amount of EU support—committed	15
Amount of national support (co-financing)—committed	5
Non-eligible cost paid by beneficiary	2
Total project value	100
Amount of EU support—paid out	15
Amount of national support (co-financing)—paid out	3
Declaration date	100
<i>Business information (ORBIS)</i>	
Name of the firm	39
Detailed address information	36
Founding year	32
NACE Rev. 2 industry (main category)	33
NACE Rev. 2 code (4 digits)	33
ORBIS size category (four categories)	38
Number of employees (last available year)	18
Volume of sales (last available year)	22
Number of companies in corporate group	17
<i>Regional characteristics</i>	
Less developed region or not (classification by NUTS-2 region)	82
Location (NUTS-3)	33
Predominantly urban region or not (classification by NUTS-3 region)	33

Total number of observations: 2,055,375. The coverage denotes the share of observations for which each variable is available. The location refers to the corresponding OP or, if the OP cannot be assigned to a region, we use the matched location information from ORBIS

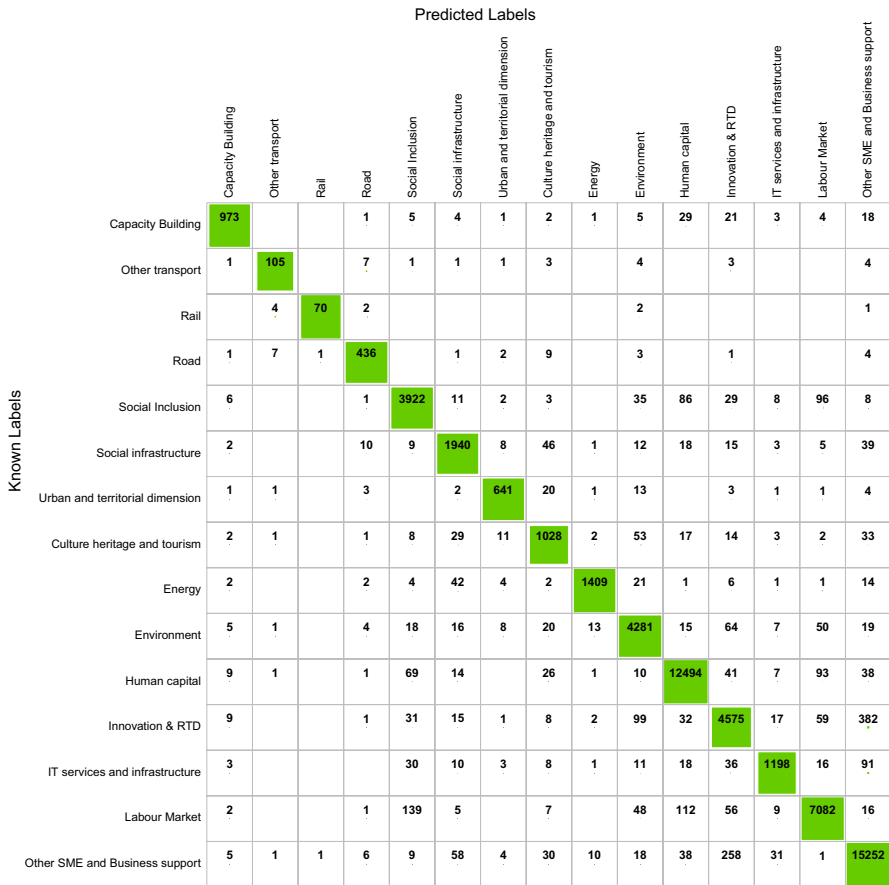


Fig. 1 Confusion matrix of the theme prediction. *Notes:* The numbers in the figure show the results of the classification of projects. The width of the green rectangles of the diagonal represents the recall and the height of the rectangles of the diagonal pictures the precision. Since all rectangles are almost filling their cell, the confusion matrix shows that the classification works well

table) for one of the ten cross folds, specifically, we choose to report the results for the cross fold with the lowest accuracy. The rows indicate the true known themes and the columns the predicted themes. Let X be the confusion matrix, then precision, recall and accuracy are defined as follows:

$$\text{precision}_j = \frac{X_{jj}}{\sum_i X_{ij}}, \quad \text{recall}_i = \frac{X_{ii}}{\sum_j X_{ij}}, \quad \text{accuracy} = \frac{\sum_i X_{ii}}{\sum_{ij} X_{ij}}. \tag{3}$$

Therefore, precision, recall and accuracy are equal to one if all the predicted themes are correct (the confusion matrix has only entries in the main diagonal) and zero if all the predicted themes are wrong (the confusion matrix has no entries in the

main diagonal). The rectangles in Fig. 1 visualise the row and column percentages of the confusion matrix. The width of the rectangles represents the row percentages and the height of the rectangles the column percentages. Therefore, the width of the rectangles of the diagonal corresponds to the recall, the height of the rectangles of the diagonal to the precision and the volumes of the rectangles of the diagonal to the squared G-measure_{*i*} ($G\text{-measure}_i = \sqrt{\text{precision}_i \times \text{recall}_i}$).

In the confusion matrix, we see that, given the true theme *Other Transport*, in 105 cases the model is able to predict the true label and in seven cases the model predicts the theme *Road*. Overall, we obtain an average classification accuracy of 0.94. However, for completeness we note that there are duplicates in the trainings and test set. Accounting for this, the average accuracy without duplicates is 0.90. In order to make the classification results easily reproducible, we assemble the **R** (R Core Team 2016) package `fastTextR` that contains an interface to the `fastText` library and is available at CRAN.²²

3.3 Comparison of the dataset with official data

We assess the validity of the assembled data by checking for outliers and plausibility and comparing its dimension with official data on regional funds assignment (equivalent to C_EU_i in Eq. 1) published by DG REGIO. Table 2 shows the sum of total project values in the database (Total value_{*i*} in Eq. 1) per country and objective, excluding projects which cannot be assigned to a specific objective. The total values in the database in general do not only consist of committed values by the EU. Multiplying the total project values with the maximum co-financing rate per country and objective (see Sect. 2.1) results in the highest amount the EU should provide. If the official committed value, given in the last column of Table 2, is lower or equal to the maximum EU's co-funding, we expect the sum of total project values in our database to be plausible. The latter is true for the large majority of member states. For Bulgaria, lists of beneficiaries are available for only less than a third of their operational programmes. The Estonian source is an online database which might not yet contain all projects. For Denmark, the gap may arise due to the total project value summing up paid-out and not committed amounts in the database.

Next, we compare the distribution of regional funds among co-financing instruments and themes as reported in the database and by DG REGIO. First, 55% of the sum of total project values correspond to projects co-funded by the ERDF (unequivocal classification) while according to data by DG REGIO around the same amount (58%) of structural funds and the Cohesion Fund is transferred via the ERDF. For ESF, the share of the total project values amounts to 22% which is exactly the same one as in official data. As several operational programmes are co-funded by the ERDF and the CF and there is no more detailed information reported in the lists of beneficiaries, we are able to attribute only about 4% of total project values in the database to the Cohesion Fund. For about 20%, we cannot clearly say which one of

²² <https://CRAN.R-project.org/package=fastTextR>.

the funds is the supporting one. Following DG REGIO data, about 20% of structural and Cohesion funds commitments are settled via the Cohesion Fund, i.e., it is likely that the major part of the not uniquely assigned total values in the lists of beneficiaries can be attributed to the Cohesion Fund.

Regarding the distribution of funds and total project values across the fifteen themes (project categories), our database proves to be consistent with official data (Fig. 2). The highest project expenditure is dedicated to *Innovation & RTD* and *Environment* according to the dataset and as also reported by DG REGIO.

4 Descriptive statistics

4.1 Stylised facts: the intraregional distribution of regional funds

In cohesion policy regulations, the EU institutions specify priority themes to be targeted by operational programmes in a programming period but do not preset any detailed requirements regarding the size or other characteristics of the projects or beneficiaries to be selected.²³ Therefore, we expect that there is no uniform strategy for distributing regional funds at the project level across European regions. The selection of projects by managing authorities is likely to depend on the accordance with an OP's underlying priority themes and its main objective, as well as the assessment of the capability of potential beneficiaries to carry out the project (with a certain volume).

One of the features of the data that has not been analysed in previous literature is the variation in average project values shown in the top right-hand side of Fig. 3. This variable ranges between EUR 15,520 per project in Marche, Italy, and EUR 57,814,068 in Ireland. While it is comparably small in Central European countries, the average project value is relatively high in some North European regions (in the UK, Denmark or Belgium) as well as the member states that joined the EU in 2004 and later. While there are additional reasons for the regional variation in the average total values such as project and beneficiary characteristics, this suggests that regions with fewer projects, on average, have higher total values per project.

The left upper part of Fig. 3 presents the number of projects per region, which varies between 32 in South-East England, UK, and over 85,000 in Puglia, Italy. There are regions with many projects like Puglia, Italy, or North Rhine-Westphalia, Germany, which are typically characterised by a high number of projects related to the themes *Labour Market* or *Human Capital*. Both themes are associated with relatively low project amounts.²⁴ Contrarily, the regions in which the largest share of funds is allocated to *Transportation Infrastructure* projects (in Poland and Croatia)

²³ It is required to take care of gender aspects and non-discrimination in the course of implementing cohesion policy.

²⁴ Refer to the Online Appendix for a descriptive analysis of the distribution of total project values among themes and other project and firm characteristics. Moreover, this finding is mirrored in estimation results in Sect. 5.

Table 2 Comparison with official DG REGIO data on regional funds (in million Euros)

Country	Obj.	Sum of total project values in database	Maximum EU co-financing rate (%)	Maximum EU co-financing	Sum of allocated EU regional funds
		Total value _i			C_EU _i
AT	1	232	75	174	102
	2	2063	50	1032	473
BE	1	1849	75	1387	451
	2	4226	50	2113	546
BG	1	2673	85	2272	6313
	2	0	85	0	0
CZ	1	22,717	85	19,309	17,555
	2	408	85	347	206
DE	1	18,445	75	13,834	10,599
	2	16,798	50	8399	4530
DK	1	0	75	0	0
	2	491	50	246	250
EE	1	1848	85	1571	2836
	2	0	85	0	0
ES	1	28,415	80	22,732	12,415
	2	10,354	50	5177	3792
FI	1	0	75	0	0
	2	6754	50	3377	998
FR	1	7300	75	5475	1876
	2	35,042	50	17,521	5289
HR	1	1592	85	1353	377
	2	0	85	0	0
IE	1	0	75	0	0
	2	6193	50	3097	414
IT	1	29,443	75	22,082	17,955
	2	16,533	50	8267	3171
LT	1	9186	85	7808	5556
	2	0	85	0	0
LU	1	0	75	0	0
	2	102	50	51	24
LV	1	7171	85	6095	3894
	2	0	85	0	0
MT	1	948	85	806	708
	2	0	85	0	0
NL	1	0	75	0	0
	2	6263	50	3132	866
PL	1	127,045	85	107,988	57,138
	2	0	85	0	0

Table 2 (continued)

Country	Obj.	Sum of total project values in database	Maximum EU co-financing rate (%)	Maximum EU co-financing	Sum of allocated EU regional funds
		Total value _i			C_EU _i
PT	1	23,880	85	20,298	10,538
	2	1912	85	1625	642
RO	1	36,888	85	31,355	16,010
	2	0	85	0	0
SE	1	0	75	0	0
	2	2571	50	1286	925
SI	1	5180	85	4403	2641
	2	0	85	0	0
SK	1	13,640	85	11,594	8144
	2	132	85	112	79
UK	1	5004	75	3753	1762
	2	9588	50	4794	3527

Source: DG REGIO, Open Data Portal for European Structural and Investment Funds

Obj. (Objective): 1 stands for *Convergence*, 2 for *Regional Competitiveness and Employment*. From our database, we do not consider funds where an unique assignment of an objective is not possible. From DG REGIO data, we do not consider commitments assigned to multiple objectives (multi objective)

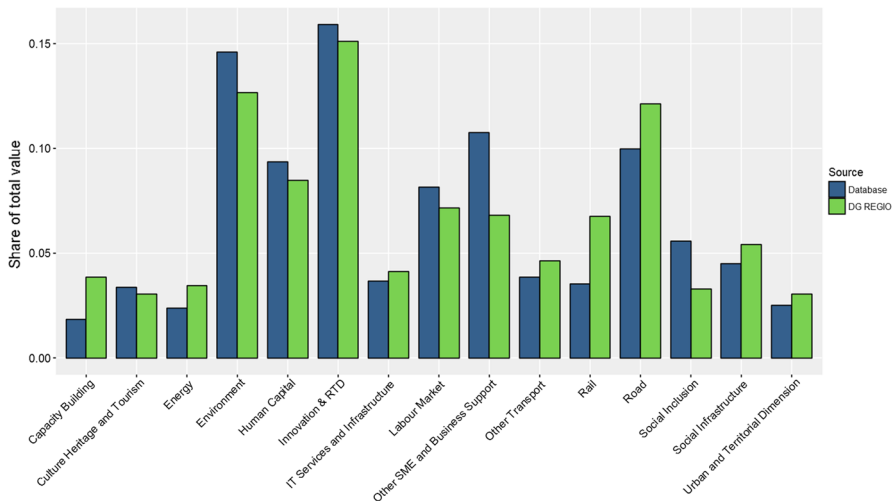


Fig. 2 Comparison of distribution of committed regional funds with DG REGIO data: shares (%) of sum of total project values. Notes: Source DG REGIO: “ERDF/ESF/CF Priority theme overview 2007–2013” downloadable at http://ec.europa.eu/regional_policy/en/policy/evaluations/data-for-research/

are characterised by relatively large project values. Also other regions, like South-East England, UK, or Vienna, Austria, have few but on average large projects. One underlying reason may be that some regions tend to report relatively many

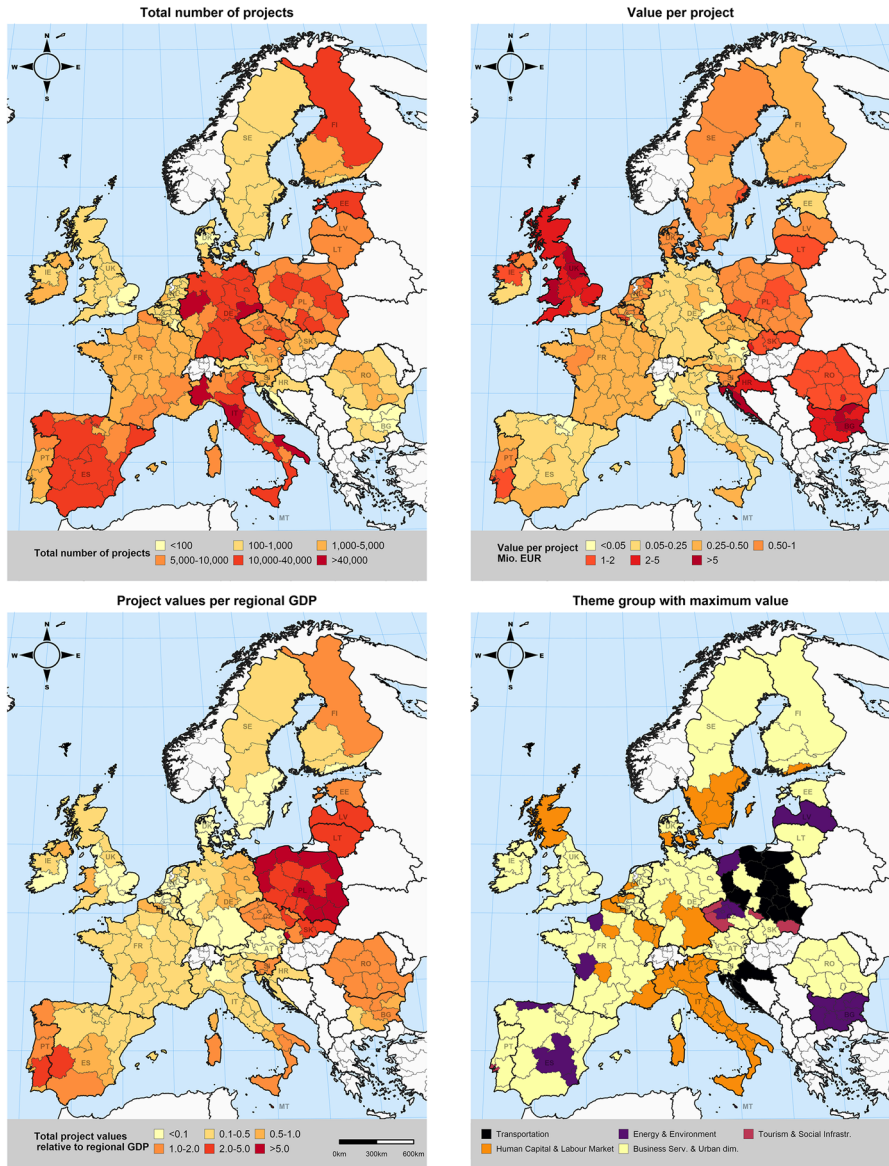


Fig. 3 Regional distribution. *Notes:* Number of projects (observations) per region: Min.: 32, 1st Qu.: 569, Median: 3220, Mean: 6526, 3rd Qu.: 8326, Max: 85,420. Average value per project (in million Euros): Min.: 0.02, 1st Qu.: 0.22, Median: 0.44, Mean: 1.18, 3rd Qu.: 1.10, Max: 14.10. Sum of regional project values as share of regional GDP (sum of 2007–2013): Min.: 0.01, 1st Qu.: 0.17, Median: 0.35, Mean: 0.91, 3rd Qu.: 1.10, Max: 8.45. Theme group with maximum sum of total project values by region: transportation refers to *Road, Rail, Other transport*; *Social inclusion* is part of the human capital group; business services are *Innovation & RTD, Other SME and business support, Capacity building, IT services and infrastructure*. Total project values are defined in Sect. 3. While the lists of beneficiaries published by the managing authorities do not permit to assign projects to regions Helsinki, Finland, as well as Dutch regions, the figure shows observations for Dutch regions that are matched with ORBIS which contains information on the location of firms

intermediate beneficiaries like public institutions on a municipal level. Those bodies apply for the funds, however, they redistribute them further or use them to carry out projects for smaller entities. In those cases, the ultimate beneficiaries are not known publicly. Finally, a part of the remaining variation in the number of projects may be largely explained by poor data availability. In Croatia, not all projects could be assigned to a NUTS-2 region given the report by the managing authority. For Bulgaria, lists of beneficiaries are only available for two out of nine operational programmes.

The dataset consists of over two million projects granted to approximately one million individual beneficiaries. That means, on average, every beneficiary receives co-financing for two projects. Only 17% of beneficiaries carry out more than one project, only 3% have more than five and only 1% more than ten. The beneficiary with the most co-financed projects (more than 18,000) is the Spanish *ICEX Espana Exportación e Inversiones*, a governmental institution that promotes (foreign) investments in Spain. The second most (more than 11,000) are carried out in the city of Florence, Italy, the third most (more than 10,000) by the governmental training and orientation section of the region of Tuscany, Italy.

Besides project and beneficiary characteristics, one stylised fact arising from the data confirms the focus of EU cohesion policy of providing most financial support to less developed regions: one main objective of the EU's cohesion policy is to support the catching-up of those regions, i.e., NUTS-2 regions with a GDP per capita below 75% of the EU-average, in economic and social terms. As one can see in Fig. 3 (left lower map), in Southern Italy, Portugal and Spain as well as Poland, Romania and Bulgaria, the Baltics or Slovenia, the total values of co-financed projects relative to the respective regional income tend to be larger than in other regions. E.g., for the Lithuania and Poland, project values account for up to eight percent of regional GDP.²⁵

Contrarily, in the so-called “Blue Banana”²⁶ the sum of total project values lies below 0.1% of regional GDP. In those regions, projects with the highest expenditure tend to be carried out by small and medium-sized firms in the education sector, in public administration as well as in professional, scientific and technical activities. Those projects are mostly aimed at *Regional Competitiveness and Employment* and the themes *Innovation & RTD* as well as *Human Capital*. The sum of project values in Scandinavian, French, Northern Italian, North-Eastern Spanish regions as well as Scotland and Northern Ireland, UK, ranges between 0.1% and 0.5% of regional GDP. Project values per capita follow a similar overall distribution. The amount in

²⁵ All of Estonia, Latvia, Lithuania, Poland, Romania, Bulgaria and Slovenia are eligible for *Convergence* (former Objective 1) funding in 2007–2013 (European Council 2006b). The region with the highest project values as percentage of its regional GDP is Podkarpackie in Southeast Poland with 8.45% or EUR 8.2 billion. According to DG REGIO, over EUR 67 billion have been committed to Poland in 2007–2013. Bulgaria and Romania joined the cohesion policy programme 2007–2013 with a lag and therefore received little funds relative to their GDP.

²⁶ Populous and usually rich regions that range from the UK, Belgium and the Netherlands, parts of Germany, Austria to Northern Italy.

most regions falls below EUR 500. The region with the highest value per capita is Bratislava, Slovakia, with EUR 12,575 per inhabitant over the course of seven years.

Table 3 presents summary statistics for the average project values in less developed versus developed regions, and confirms this finding. Projects carried out by beneficiaries in less developed regions are larger on average. A *t*-test on the significance of the mean differences shows that they are statistically significant at the 1%-level. Most of the projects there correspond to the *Road, Environment* but also *Innovation & RTD* themes and are carried out by very large firms that operate in public administration and the manufacturing industry. By contrast, large shares of regional funds in higher-income regions go to *Innovation & RTD, Environment* and *Labour market* projects, while less money is dedicated to transportation infrastructure.

Gagliardi and Percoco (2017) indicate that the effectiveness of EU cohesion policy varies across urban and rural (NUTS-3) regions, thus, we are also interested in potential differences in the usage of regional funds in such areas. The summary statistics of project values shown in Table 3 reveal statistically significantly higher average single project values in (predominantly) urban NUTS-3 regions. The following section provides more details on the differences in regional funds distribution within urban versus within rural regions.

4.2 The intraregional distribution of funds with respect to project and firm characteristics

Type of fund, objectives and themes

The allocation of regional funds to OPs co-financed by different types of funds and under different main objectives is closely related to previous findings regarding the development status of regions and the choice of thematic priorities.²⁷ The total values of projects subsidised by the ERDF sum to roughly EUR 270 billion compared to EUR 107 billion from the ESF and EUR 18 billion from the CF. In addition, projects accounting for EUR 100 billion are most likely funded by the CF. Typically, the Cohesion Fund (CF), for which only regions in countries with a gross national income below 90% of the EU average are eligible, is targeted at co-financing (large) infrastructure projects and the accessibility of lagging-behind regions. A major part of the European Social Fund (ESF) is allocated to projects with relatively smaller project values fostering, e.g., human capital, while the ERDF co-finances a broad spectrum of project types. These funding priorities are mirrored in the volume of corresponding projects: the median CF project value amounts to about EUR 100,000 compared to EUR 37,000 for the ERDF and EUR 3200 for the ESF.

Co-financing by different funds additionally varies across types of regions (Table 4). By construction, the CF is more important in less developed regions, while the ESF accounts for only around 10% of project volumes in those poorest regions. It is interesting that the CF plays a larger role for co-funded project volumes

²⁷ This allocation is negotiated when setting up the OPs, before the actual process of project selection.

Table 3 Summary statistics in EUR 1000

	Min	Median	Mean	Max
Average project value, developed regions	15	352	514	4226
Average project value, less developed regions	89	848	1392	9032
Average project value, urban regions	0	242	1167	29,962
Average project value, rural regions	2	150	641	19,530

Minimum of zero is due to rounding. Some projects have project values below EUR 1000

in (predominantly) urban regions as compared to rural ones. According to the data, this may be related to the fact that there are more relatively large *Road* projects funded in areas with higher population density.

The overall objective of the OP each project is part of, is closely related to the development status of the regions, as, along with several exemptions, only less developed regions are eligible for *Convergence* funds. While around half of the (number of) projects in our sample is aimed at *Convergence*, their value is three times as large as that of *Regional Competitiveness and Employment* projects. As expected, operational programmes in less developed regions have a focus on *Convergence*, while OPs in richer regions receive more to boost *Regional Competitiveness and Employment*.

For the appropriate targeting of regional funds, the European Commission defines fifteen priority themes (see Sect. 3.2) to classify projects. In total, the largest sums are committed to projects in categories *Innovation & RTD* (EUR 70 billion), *Environment* (EUR 65 billion) and *Other SME and Business Support* (EUR 46 billion), followed by *Labour Market*, *Road* and *Human Capital* projects. The number of observations per theme ranges from 1058 *Rail* projects to over 600,000 projects related to the *Labour Market*. There is also considerable variation of individual project values across themes. Projects related to *Transportation* (EUR 1–3 million), *Urban and Territorial Dimension* as well as *Culture Heritage and Tourism* (EUR 200,000 each) are the largest. *Human Capital* (EUR 6000), *Labour Market* (EUR 2200) and *Energy* (EUR 1250) projects are the smallest.

Additionally, the lower right part of Fig. 3 shows the theme at which the maximum sum of project values in a region is targeted. For better readability the map shows five groups of themes instead of all fifteen.²⁸ In total, the largest sums in Poland and Croatia are related to *Transportation* projects. *Energy* and *Environment* projects are most important in Latvia, parts of the Czech Republic, two French regions, three Spanish regions and a Polish region. Contributions to *Culture*, *Tourism* and *Social Infrastructure* are most pronounced in some Czech regions and East Slovakia. Scotland, South Sweden, almost all of Italy and parts of Germany and France have their largest project sums related to *Human Capital*, the *Labour Market*

²⁸ The aggregation of the themes is based on a hierarchical clustering, where we use the cosine dissimilarity as distance measure and Ward's method (Ward Jr 1963) for the clustering. More information about the clustering of text data can be found in Feinerer et al. (2008).

Table 4 Summary statistics by type of fund, theme and beneficiary size

Category	Share of regional sum of project values				
	Total sample (%)	Less developed regions (%)	Developed regions (%)	Rural regions (%)	Urban regions (%)
CF	4	4	3	5	3
ERDF	55	56	60	62	42
ERDF or CF	20	30	7	18	42
ESF	22	10	30	14	13
Capacity building	2	2	2	1	1
Culture heritage and tourism	3	4	2	5	2
Energy	2	3	3	3	2
Environment	15	15	15	24	10
Human capital	9	5	8	6	6
Innovation & RTD	16	15	20	20	16
IT services and infrastructure	4	2	9	3	4
Labour market	8	4	15	6	5
Other SME and business support	11	11	6	15	9
Other transport	4	3	4	3	6
Rail	4	5	1	0	9
Road	10	17	1	2	23
Social inclusion	6	4	10	4	3
Social infrastructure	5	6	3	7	3
Urban and territorial dimension	3	4	1	1	0
Large company	17	20	13	26	15
Medium sized company	17	20	12	23	14
Small company	39	28	53	37	27
Very large company	27	32	22	14	45

Summary statistics of total project values by type of funds (funding instrument), theme and beneficiaries' size respectively in thousand Euros. Sums of total values for first subset: 495 billion EUR—238 bill. EUR (less dev.), 121 bill. EUR (dev.)—87 bill. EUR (rural), 121 bill. EUR (urban), second subset: 442 bill. EUR—223 bill. EUR (less dev.), 101 bill. (dev.)—81 bill. EUR (rural), 109 bill. (urban), third subset: 305 bill. EUR—179 bill. EUR (less dev.), 84 bill. EUR (developed)—87 bill. EUR (rural), 121 bill. EUR (urban). Shares of total value may not add up to 100 due to rounding. Total number of observations for first subset: 2,039,648–479,032 (less dev.), 589,451 (developed)—310,350 (rural), 223,746 (urban); second subset: 1,692,990–372,522 (less dev.), 416,879 (developed)—310,350 (rural), 188,237 (urban); third subset: 778,196–310,754 (less dev.), 335,997 (developed)—310,350 (rural), 223,746 (urban). For details regarding the data coverage of each variable refer to the previous section

and *Social Inclusion*. The largest project sums in the rest of Europe are associated with the fifth category that includes *SME and Business Services, IT services and infrastructure, Capacity Building, Innovation & RTD* as well as *Urban and Territorial Dimension*.

Comparing all less developed to all higher-income regions, differences in the average distribution of total project values across priority themes become apparent

(see Table 4). For example, *Road* accounts for 17% of the sum of project values in less developed regions and only for 1% in higher-income regions. Moreover, the share targeted at *Labour Market* projects in richer regions exceeds that in poorer ones by 11 percentage points.

Industrial structure and firm size

The matched data from ORBIS allows us to go into more detail with respect to characteristics of the beneficiaries. Figure 4 shows the distribution of funds across NACE Rev. 2 industries. The left-hand side illustrates total sums while the right-hand side shows the distribution of individual project values. The matching process and the coverage in ORBIS enable us to assign an industry to one third of observations, i.e., almost 700,000 projects.

Overall, Fig. 4 (left part) shows that projects with the highest sum of total values are carried out by firms or institutions operating in public administration, defence and social security (EUR 80 billion), transportation and manufacturing (EUR 30 billion each) and education (EUR 20 billion). The right-hand side of Fig. 4 shows that the single project volumes also vary across beneficiaries in different industries. While the median value of projects carried out by public beneficiaries (NACE Rev. 2 industry O) lies clearly below EUR 100,000, firms operating in the energy and water sector are responsible for the projects with highest median values (over EUR 200,000). Most of the latter projects have to do with electricity production, air conditioning, water supply as well as waste collection and treatment.

Beneficiary firms or institutions also differ in their size. Table 4 shows the distribution of regional funds across four categories of firm size defined by ORBIS.²⁹ In terms of overall project sums, small firms make up the largest share (EUR 120 billion), followed by very large firms (EUR 80 billion) and medium-sized as well as large firms (EUR 50 billion each). Note that there are fifteen times more small beneficiaries of structural funds than very large ones. They are especially strongly represented in projects co-financed by the ESF. In total, 85% of recipients in the database are small or medium-sized (SME) companies, which reflects the priorities set out in the Community's strategic guidelines on cohesion. Moreover, this may indicate that the projects, which smaller and medium-sized firms submitted in order to apply for co-funding and which were selected, are smaller in terms of single project values than the projects carried out by large firms.

Interestingly, Table 4 shows that while managing authorities in regions with an income above 75% of the EU average (developed ones) select mostly projects carried out by small firms, in less developed regions the volumes of projects carried out by very large companies is higher. Also the share of project values administered by SMEs in less developed regions falls behind the share in higher-income regions by 17 percentage points. There is a similar heterogeneity when comparing projects in rural and urban regions, and it is the share of the value of projects carried out by very large firms in urban regions which

²⁹ ORBIS considers companies to be small if they could not be classified, which could possibly inflate the number of small companies due to missing data. However, labelling the observations based on the number of employees, we find a similar distribution of labels.

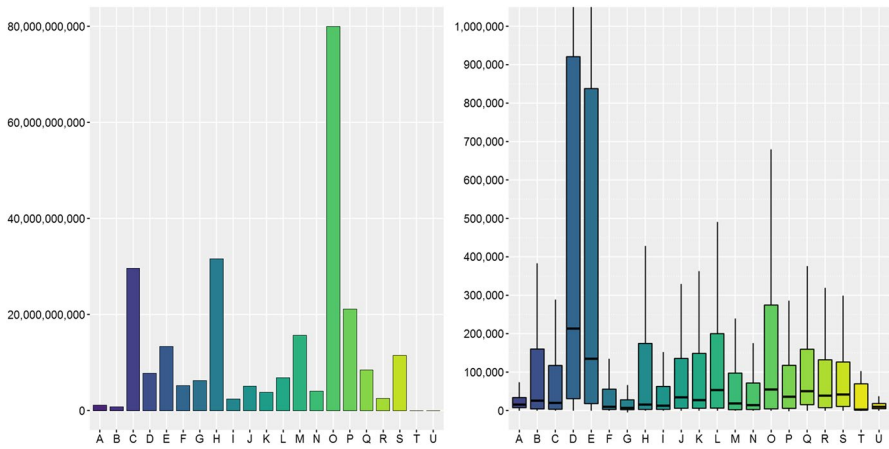


Fig. 4 Sum of total project values and single project values by NACE Rev. 2 industry. *Notes:* Left: sum of total project values by NACE Rev. 2 industry. Right: distribution of values per project; dark horizontal line marks the median. A: agriculture, forestry, fishing; B: mining, quarrying; C: manufacturing; D: energy; E: water supply, sewerage, waste management; F: construction; G: wholesale; H: transportation; I: accommodation and food services; J: information and communication; K: financial services; L: real estate services; M: professional, scientific and technical activities; N: administration and support activities; O: public administration, defence, social security; P: education; Q: human health and social work; R: arts, entertainment, recreation; S: other services, T: household services, U: activities of extraterritorial organisations. The figure represents 33% of observations and 51% of the sum of total values

exceeds the one in rural areas by more than 30 percentage points. This pattern may be driven by the fact that cities implement relatively many large projects.

5 Focus on the single beneficiary

Given the evidence on substantial variation in the total value of single projects, we investigate the potential determinants of this variable for the projects in our database. In particular we hypothesise that the project size is closely related to the following characteristics: (1) the funding instrument (ERDF, ESF or CF), as certain types of funds typically co-finance different kinds of projects, (2) the objective of the corresponding OP (*Convergence or Regional Competitiveness and Employment*), which mirrors the development status of the region, (3) the managing authorities’ focus on thematic priorities (themes), e.g., a project which consists of building road infrastructure will have a higher value than a specific employee’s training, and (4) beneficiaries’ characteristics like size or industry. Regarding firm size, our hypothesis is that, on average, larger firms are capable of carrying out larger projects than smaller firms.

To this end, we estimate the following equation:

$$\ln(TV_i) = \alpha + \psi R_i + \gamma F_i + \delta O_i + \rho T_i + \phi I_i + \omega S_i + \varepsilon \tag{4}$$

where $\ln(TV_i)$ represents the logarithm of the total project value of project i , R_i the region in which project i is located (a sort of regional fixed effect), F_i the type of

fund which supports project i , O_i the objective of the funding for project i , T_i the theme under which project i is supported, I_i the industry of the beneficiary and S_i the size of the beneficiary of project i . The variables R_i , F_i , O_i , T_i , I_i and S_i are factor variables and the category with the median coefficient is always the one that is excluded respectively (in the following Tables). Thus, the resulting coefficients should be interpreted relative to the project with the coefficient being in the middle of the (conditional) distribution of this variable, which is always shown in the results tables with a coefficient of zero. Significance levels are not shown, as they depend entirely on the chosen reference category and have no general meaning in such a context.³⁰ We do not include projects with total values which are zero or negative and end up with 482,040 observations for which Eq. (4) is estimated.³¹

In this way, we are able to shed light on conditional differences between the volume of projects supported by different funds with different objectives and themes as well as with beneficiaries in different industries and of different size. Moreover, we are able to analyse whether unexplained residual variation in project volumes (in countries and regions) exists. We think that this could entail interesting policy implications as Locatelli et al. (2017) show for public procurement processes that funding larger projects prepares the ground for more corruption than it would be the case when more and smaller projects are tendered. Other analyses show that firm subsidies are not equally effective when granted to firms of different size (Criscuolo et al. 2012) and to firms located in different regions (Bachtrögler et al. 2018).

In order to take the heterogeneity of funding principles into account, we analyse total values of single projects not only for the complete sample but additionally run Eq. (4) separately for, first, projects co-funded by (1) the ERDF and the ESF (structural funds) and (2) the CF. Second, we split the sample into projects carried out by (1) public and (2) private beneficiaries.³² Finally, in a robustness check in which we model firm size by the number of employees, volume of sales and add firm age as a further control variable, we run regressions using sub-samples of projects carried out in (1) less developed and (2) developed regions, as well as, (3) urban and (4) rural regions.

5.1 Type of fund, objective and themes

Controlling for the other variables included in Eq. (4), the projects with the smallest total value can be identified as co-funded by the ESF (left panel of Table 5).

³⁰ While a coefficient might be highly significant with respect to reference category A, it could be not significant with respect to category B. Thus, one might create any significance level by choosing a respective reference category. Only the coefficient keeps its meaning. Wherever there is a specific meaning to the significance level, such as for revenues or the number of employees of a beneficiary, we also report the significance levels.

³¹ Only observations matched with ORBIS data can be considered for the econometric exercises in this section.

³² Here, we consider beneficiaries in the NACE Rev. 2 sector “O—Public administration and defense; compulsory social security” (according to ORBIS data) as public firms, whereas all other sectors are categorised as private.

They are more than three-quarters smaller than those co-funded by the ERDF and (on average) by the CF (if the projects that cannot be assigned clearly to the ERDF or the CF are counted as CF co-funded ones). This corresponds to the goals of the various funds, as the ESF is mostly funding smaller projects related to the inclusion, training and adaptability of workers in the labour market and employment, while the ERDF and CF funds are aimed at improving the economic structure and fundamentals of regions (European Commission 2017). Splitting the sample by broad industries reveals that CF projects carried out by public entities are on average twice as large in terms of their project value than ESF projects. Likewise, relative to ESF project values, CF projects are (on average) largest in the case of beneficiaries in the private sector.

Furthermore, controlling for everything else, Table 5 shows that projects corresponding to OPs with the *Convergence* objective are larger than projects under the *Regional Competitiveness and Employment* objective (right panel of Table 5).

Turning to another aspect, large differences in conditional project values arise also with respect to the projects' themes. 42% of the largest projects in the database (with a volume above EUR 50 million) are associated with the three transport-related themes (*Rail, Road, Other Transport*). As also found in the unconditional analysis in Sect. 4, it might not be surprising that the projects with themes related to *Labour Market, Human Capital and SME* are among the projects with the lowest conditional project values, but the *Energy* theme would probably be expected to be among the one with the highest project values (see Table 6). Regarding the latter theme, there is large variation across subsets and it turns out that especially *Energy* projects co-financed by the CF and carried out by private-sector (but not public-sector) beneficiaries are rather small.

In general, the largest projects co-financed across all subsets are targeted at *Road and Other Transportation* infrastructure (and fostering the *Urban and Territorial Dimension* in the case of the CF). Interestingly, relatively bigger investment projects in *Rail* systems appear to be co-financed by structural funds (and not the CF) and implemented by private entities (not public ones). The same is true for projects with a relatively high conditional value in the fields of *Social Infrastructure, Environment and Innovation & RTD*.³³

Moreover, conditional differences in project values across themes do not necessarily correspond to the differences across industries. An interesting observation is that the lowest project values can be observed for the *Energy* theme, but the highest conditional project values for the energy industry (Sect. 5.2). The maximum project value within the *Energy* theme as well as the energy industry is found for the same project which amounts to EUR 906 million. However, the *Energy* theme includes 7715 projects with a project value of EUR 1250 each and 64% of the projects have a value below EUR 5000. Those projects below EUR 5000 are not conducted by firms in the energy industry but are observed mainly in "Real Estate activities" (59%), "Manufacturing" (11%) and "Wholesale/retail; repair vehicles" (8%).

³³ Note that the education sector (NACE Rev. 2) includes both private and public entities but is considered as belonging to the private sector in this analysis.

Table 5 Regression coefficients with respect to type of fund and objective

Type of fund	Coef. (Eq. 4)			Objective (Eq. 4)	Coef.		
	General (%)	Public (%)	Private (%)		General (%)	Public (%)	Private (%)
ERDF or CF	+ 129	+ 158	+ 98	Convergence	+18	+ 17	+ 28
ERDF	+ 78	+ 25	+ 61	Regional Comp. and Employment	0	0	0
CF	+ 35	+ 41	+ 53				
ESF	0	0	0				

Coefficients of regressions based on Eq. (4). White-robust standard errors. Dependent variable: logarithm of the total project value. Other control variables (region, industry, theme, firm size) are also included but reported in previous and subsequent tables. Coefficients of respective dummy variable relative to the median category. *ERDF* European Regional Development Fund, *ESF* European Social Fund, *CF* Cohesion Fund. Columns “General”: 482,040 degrees of freedom, adj. R^2 of 0.46. Columns “Public”: 69,351; Adj. R^2 0.62. Columns “Private”: 412,689; Adj. R^2 0.48

5.2 Industry and beneficiaries' characteristics

The highest conditional project values are attributed to beneficiaries in the *Energy* and *Water* industries (Table 7). The largest three project values of firms in the *Energy* industry are EUR 906 million, EUR 103 million and EUR 100 million and only 37% of the projects have a total value below EUR 100,000. Within the *Wholesale* industry, on the other hand, only 12% of the projects have total values larger than EUR 100,000.

As expected, firm size also plays a role for the level of the total project value (Table 8). Overall, we see that the larger the beneficiary, the larger is the total value. However, the difference in project volumes is surprisingly low, as the volume of a project of a very large company is not even twice the one of a medium-sized company but the average revenue of the very large companies in our sample (EUR 653 million) is approximately 130 times the average revenue of the medium-sized companies (EUR 5 million). The pattern is similar when considering the samples split according to the co-financing type of fund and the sector the beneficiary operates in.

5.3 Residual variation in single project values

Comparing the total value of single projects across NUTS-2 regions, while controlling for the type of fund, objective, theme, industry and size of the beneficiary, reveals differences of more than plus and minus 400%. The twenty top and bottom regions are presented in Table 9. The lowest conditional amounts per project can be found in Austria, Spain, Estonia, Germany and Belgium. The highest conditional values per project can be observed in the UK, the Netherlands, Finland, Malta and

Table 6 Regression coefficients with respect to theme

Theme	General (Eq. 4)	Structural funds (Eq. 4)	Cohesion fund (Eq. 4)	Public (Eq. 4)	Private (Eq. 4)
Road	+ 270%	+ 238%	+ 225%	+ 248%	+ 151%
Other transport	+ 216%	+ 191%	+ 85%	+ 143%	+ 218%
Rail	+ 186%	+ 124%	+ 11%	- 94%	+ 208%
Culture heritage and tourism	+ 69%	+ 73%	- 32%	+ 59%	+ 70%
Urban and territorial dimension	+ 45%	+ 46%	+ 99%	- 12%	+ 57%
Social infrastructure	+ 40%	+ 37%	- 62%	- 8%	+ 94%
Environment	+ 29%	+ 37%	- 128%	- 18%	+ 49%
Innovation & RTD	+ 29%	+ 40%	- 100%	- 4%	+ 52%
IT services and infrastructure	0%	0%	0%	0%	0%
Labour market	- 14%	- 38%	- 81%	- 43%	+ 11%
Capacity building	- 15%	- 22%	- 143%	- 44%	+ 3%
Social inclusion	- 22%	- 48%	- 130%	- 87%	+ 12%
Other SME and business support	- 66%	- 67%	- 129%	- 353%	- 27%
Human capital	- 68%	- 94%	- 3%	- 65%	- 27%
Energy	- 124%	- 15%	- 296%	61%	- 92%
Adj. R ²	0.49	0.46	0.70	0.62	0.48
No. of observations	482,040	452,726	29,314	69,351	412,689

Coefficients of regressions based on Eq. (4). White-robust standard errors. Dependent variable: logarithm of the total project value. Other control variables (firm size, industry, region, fund type and objective) are also included but reported in other tables. Coefficients of respective dummy variable relative to the median category. “Type of fund” is excluded as control in columns 3 and 4. “NACE industry” is excluded as control in columns 5 and 6

Luxembourg.³⁴ Beneficiaries which are similar in size, are in the same industry, receive money from the same fund and within the same theme, but are located in Lower Austria (NUTS-2 region AT11) on average have about 8.5 times lower project values than beneficiaries located in the East of England (NUTS-1 region UKH). Apparently, in Lower Austria, 90% of projects are smaller than EUR 10,000 and 4688 projects are smaller than EUR 1000, whereas in the East of England the smallest project is already EUR 133,168 in size.

Table 10 shows regional effects based on estimating Eq. (4) for subsets related to project and beneficiary characteristics. First, the upper panel shows the ten regions in which, conditional on the other control variables, the biggest and smallest projects co-funded by structural funds on the one hand, and the Cohesion Fund on the

³⁴ Interestingly, the top six UK regions had a majority vote for “Leave” in the Brexit referendum 2016.

Table 7 Regression coefficients with respect to NACE-industry

Industry NACE code	General (Eq. 4)	Structural funds (Eq. 4)	Cohesion fund (Eq. 4)
D—Energy (without water)	+ 145%	+ 112%	+ 177%
E—Water supply; sewerage, waste mgmt. and rem. act.	+ 89%	+ 44%	+ 156%
Q—Human health and social work activities	+ 52%	+ 50%	+ 19%
P—Education	+ 31%	+ 25%	+ 92%
T—Activities of households as employers	+ 29%	+ 44%	− 157%
U—Activities of extraterr. organisations and bodies	+ 23%	+ 14%	−
B—Mining and quarrying	+ 21%	+ 23%	+ 68%
M—Professional, scientific and techn. activities	+ 15%	+ 14%	+ 20%
J—Information and communication	+ 13%	+ 17%	+ 20%
S—Other service activities	+ 10%	+ 6%	− 30%
H—Transportation and storage	+ 5%	− 9%	+ 69%
O—Public adm. and defense; compulsory social security	+ 3%	+ 5%	+ 24%
R—Arts, entertainment and recreation	0%	0%	0%
C—Manufacturing	− 2%	+ 3%	+ 17%
I—Accommodation and food service activities	− 12%	− 15%	+ 17%
L—Real estate activities	− 14%	+ 47%	− 101%
N—Administrative and support service activities	− 15%	− 14%	− 76%
K—Financial and insurance activities	− 20%	− 18%	− 57%
F—Construction	− 21%	− 23%	− 1%
G—Wholesale/retail; repair vehicles	− 41%	− 43%	− 8%
A—Agriculture, forestry and fishing	− 43%	− 37%	− 55%
Adj. R ²	0.49	0.46	0.70
No. of observations	482,040	452,726	29,314

Coefficients of regressions based on Eq. (4). White-robust standard errors. Dependent variable: logarithm of the total project value. Other control variables (firm size, region, fund type and objective, theme) are also included but reported in previous and subsequent tables. Coefficients of respective dummy variable relative to the median category. “Type of fund” is excluded as control in columns 3 and 4

Table 8 Regression coefficients with respect to the beneficiaries’ size

Theme	General (Eq. 4)	Structural funds (Eq. 4)	Cohesion fund (Eq. 4)	Public (Eq. 4)	Private (Eq. 4)
Very large company	+ 91%	+ 84%	+ 131%	+ 91%	+ 109%
Large company	+ 32%	+ 31%	+ 44%	+ 32%	+ 37%
Medium-sized company	0 %	0%	0%	0%	0%
Small company	− 31%	− 31%	− 29%	− 30%	− 27%
Adj. R ²	0.49	0.46	0.70	0.62	0.48
No. of observations	482,040	452,726	29,314	69,351	412,689

Coefficients of regressions based on Eq. (4). Dependent variable: logarithm of the total project value. Other control variables (firm size, industry, region, fund type and objective) are also included but reported in other tables. Coefficients of respective dummy variable relative to the median category

Table 9 Regression coefficients with respect to NUTS-2 region

Top 20	(Eq. 4)		Bottom 20	(Eq. 4)			
UKH	+ 405%	MT00	+ 234%	ES52	- 254%	ES42	- 279%
UKE	+ 372%	UKM	+ 224%	DE1	- 255%	ES12	- 297%
UKL	+ 338%	NL12	+ 223%	BE35	- 256%	ES53	- 309%
UKC	+ 314%	NL31	+ 222%	DE6	- 258%	ES22	- 310%
NL13	+ 296%	NL33	+ 219%	BE34	- 260%	ES41	- 310%
NL11	+ 266%	NL32	+ 214%	ES24	- 262%	EE00	- 318%
UKG	+ 259%	UKD	+ 198%	DEA	- 263%	ES23	- 349%
UKK	+ 253%	NL21	+ 192%	AT13	- 265%	ES13	- 362%
NL22	+ 236%	NL23	+ 190%	ES21	- 268%	AT11	- 376%
FI1B	+ 234%	LU00	+ 188%	ES62	- 269%	AT12	- 461%

Coefficients of regressions based on Eq. (4) (482,040 degrees of freedom, adj. R^2 of 0.49). White-robust standard errors. Dependent variable: logarithm of the total project value. Other control variables (firm size, industry, fund type and objective, theme) are also included but reported in subsequent tables. Coefficients are relative to the median NUTS2-region which is FR25

other hand, were carried out. While the results for the first case broadly reflect the regional effects identified for the complete sample, the Cohesion Fund sample only includes a limited set of countries. From the regions located in those lagging-behind countries, projects in Estonia, the Czech Republic as well as some parts of Spain are smallest.

The lower panel of Table 10 indicates that there may be differences in selecting private and public beneficiaries across countries. While Austrian regions and Estonia form part of the Bottom 10-group only in the case of private firms, the smallest conditional project values for beneficiaries in the public sector are present in Spanish, one Italian and two German regions. Interestingly, next to British and Dutch regions, projects by public institutions are relatively big in size in Helsinki-Uusima (FI1B), Finland, and those of firms in the private sector in Malta.

The estimation of Eq. (4) controls for regional fixed effects, thus, any regional influence on the remaining differences between total project values of single projects should have been removed. Figure 5 shows the average residuals of a regression similar to Eq. (4), but excluding the regional control variables (R_i). Hence, the figure shows the (average) unexplained part of the project value when controlling for fund type, objective, theme, industry and firm size split by country. It confirms that the projects, which are similar in many dimensions but the region, with the lowest total values can be found in Austria, Estonia and Spain and the ones with the highest values in Luxembourg, the Netherlands, Denmark and the UK.

Due to the findings in Sect. 4 and as Luxembourg, the Netherlands, the UK and Denmark are relatively densely populated, one could suggest that population density determines that finding. However, the relationship is not that clear. The residual variation for Finland, which is the least densely populated country in Europe, is also relatively high, while Germany, which is relatively densely populated, shows a downward bias with regard to conditional average project

Table 10 NUTS-2 regression coefficients for subsamples

Structural funds				Cohesion fund			
Top 10		Bottom 10		Top 10		Bottom 10	
UKH	+ 396%	DEA	- 277%	ES70	+ 469%	CZ03	- 193%
UKE	+ 390%	DE1	- 278%	HR03	+ 254%	CZ05	- 193%
UKL	+ 347%	ES53	- 291%	ES21	+ 245%	CZ04	- 198%
UKC	+ 333%	ES22	- 295%	HR04	+ 192%	CZ06	- 206%
NL13	+ 316%	ES12	- 296%	RO32	+ 177%	CZ08	- 207%
NL11	+ 287%	ES41	- 306%	PT20	+ 171%	CZ07	- 207%
UKG	+ 281%	ES23	- 332%	RO31	+ 155%	CZ01	- 213%
UKK	+ 276%	ES13	- 352%	RO11	+ 147%	ES	- 214%
NL33	+ 258%	AT11	- 381%	RO21	+ 137%	ES64	- 230%
NL31	+ 258%	AT12	- 451%	PT30	+ 123%	EE00	- 404%
Public beneficiaries				Private beneficiaries			
Top 10		Bottom 10		Top 10		Bottom 10	
UKE	+ 544%	ES43	- 199%	UKH	+ 371%	ES13	- 264%
UKC	+ 468%	ES64	- 228%	UKL	+ 361%	ES51	- 267%
UKL	+ 394%	ES22	- 233%	UKE	+ 345%	AT22	- 270%
FI1B	+ 374%	ES12	- 235%	MT00	+ 316%	ES41	- 272%
UKM	+ 332%	ES23	- 236%	UKC	+ 297%	ES53	- 278%
UKH	+ 332%	DE6	- 264%	UKK	+ 278%	AT13	- 281%
UKG	+ 329%	ES53	- 267%	NL11	+ 246%	ES23	- 306%
NL12	+ 327%	DEB	- 281%	UKG	+ 246%	EE00	- 308%
NL13	+ 322%	ITC4	- 304%	NL13	+ 235%	AT11	- 375%
NL23	+ 303%	ES13	- 351%	UKM	+ 222%	AT12	- 492%

Dependent variable: logarithm of the total project value. Coefficients are relative to the median NUTS2-region which is FR25. White-robust standard errors. Upper panel, columns 1–4: Subsample with only projects co-funded by structural funds (ERDF and ESF). Number of observations: 452,726, Adjusted R^2 : 0.46. Upper panel, columns 5–8: Subsample with only CF projects. Number of observations: 29,314, Adjusted R^2 : 0.70. Lower panel, columns 1–4: Beneficiaries in the public sector (NACE Rev. 2 sector O). Number of observations: 69,351, Adjusted R^2 : 0.63. Lower panel, columns 5–8: Beneficiaries in the private sector and not NACE Rev. 2 sector O, respectively. Number of observations: 412,689, Adjusted R^2 : 0.48

volumes. Referring to Locatelli et al. (2017), the countries with the largest residuals are characterised by relatively low levels of corruption and good institutions at the regional level (Charron et al. 2015). Therefore, we suspect that the variation may be determined by historically grown funding strategies and further national institutional settings. Verifying this assumption remains for future research.

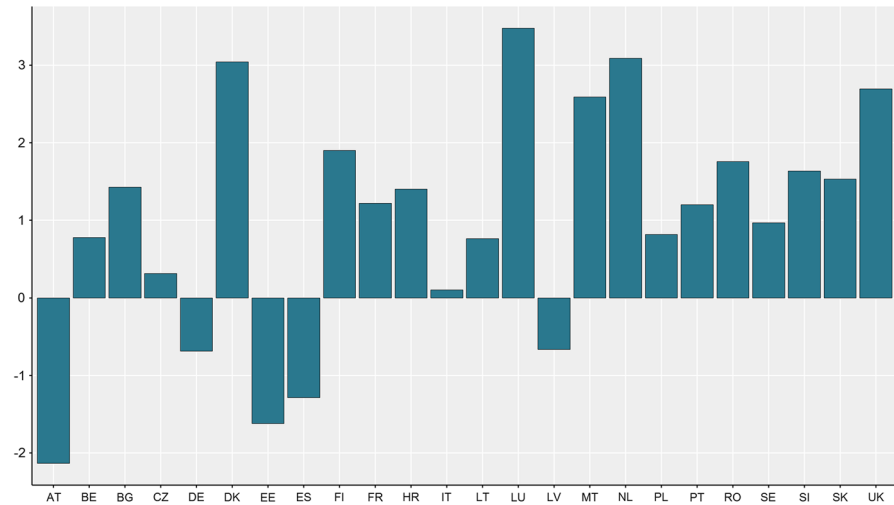


Fig. 5 Comparison of residuals of regression of Eq. (4) by country

6 Conclusion

The novel database introduced in this study contains detailed information on over two million projects co-financed by the European Regional Development Fund, the European Social Fund and the Cohesion Fund in 25 EU member states in the programming period 2007–2013. Additional to project information such as the total value of each project and a project category (theme), the beneficiaries are matched with the ORBIS business database.

This study shows that there are different patterns in the intraregional funds distribution across and within countries, both in terms of project and beneficiary characteristics. Moreover, the analysis points to the fact that managing authorities select different kinds of projects, with significantly different single project volumes, in urban and rural regions. The same turns out to be true for less developed and other NUTS-2 regions, which seems to be linked to the priorities and regulations of certain types of funds, e.g., the CF, and main objectives.

In addition, we find that most regional funds are dedicated to transportation infrastructure in less developed regions, whereas in higher-income regions a larger focus is put on fostering *Labour market* and *Social inclusion*. In all regions, *Innovation and RTD* as well as *Environment* projects form a large share of the sum of project values. Regarding beneficiaries' characteristics, the largest share—on average around 40%—of the (unconditional) sum is allocated to projects carried out by small firms, whereas this does not hold true for less developed and urban regions.

In the econometric analysis we test for the importance of certain project and beneficiaries' characteristics in determining a single project's size. The largest single projects in terms of their total value are co-funded by the ERDF and the CF (as compared to the ESF), and under the *Convergence* objective (as compared to *Regional Competitiveness and Employment*). In line with the priorities of the different funding instruments, the largest

projects are attributed to transportation infrastructure projects (*Road, Rail and Other Transport*). Regarding the beneficiary, larger firms (with higher revenues and more employees) carry out projects with higher total value, however, the average single project value of (very) large firms is only about twice as high as that of small entities. Having controlled for all characteristics, some variation in single project values remains unexplained, and this residual variation appears to differ across countries. From that we draw the conclusion that national institutional settings or traditional funding procedures may play a role.

We contribute to the academic and political debate by making a dimension of EU cohesion policy implementation visible that has not gained much attention until now. The possibility to compare individual projects' and beneficiaries' characteristics across heterogeneous regions and countries opens a new strand of research questions, e.g., whether projects in a region are carried out by firms located in the same region and by what this is determined. Moreover, the data could feed in dynamic stochastic general equilibrium or other forecasting models that simulate potential policy outcomes under different scenarios. Finally, the analysis of this dataset may entail interesting conclusions on more or less effective ways of distributing regional funds within European regions. In this respect, we also think that future research could explore how the considerable residual variation in the size of projects found in this paper is related to institutional settings in the respective countries as well as to the efficiency and effectiveness of EU cohesion policy.

Acknowledgements Open access funding provided by Vienna University of Economics and Business (WU). Project funding by the Vienna University of Economics and Business (Institute for International Economics and "small-scale project by WU junior faculty") is greatly appreciated. The authors gratefully acknowledge assistance by the following students: Konrad Fätkenheuer, Maximilian Fritz, Maria Kodajova, Christian Lubi, Andrej Mijakovic, Anna Stelzer and Tinja Zerzer. Moreover, we appreciate very helpful comments and suggestions by Peter Huber and two anonymous reviewers. Moreover, we thank participants at the WU Department of Economics seminar, the Geoffrey Hewings Regional Economics Workshop and a workshop at the Joint Research Centre in 2016 for fruitful discussions. The opinions expressed in this paper are those of the authors only and do not necessarily reflect the official viewpoint of the German Council of Economic Experts.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

Appendix

List of operational programmes and coverage in the novel database

In the following table, we first report the fund co-financing the particular OP and the objective under which it operates as well as the geographical unit concerned (NUTS-2 or NUTS-1 or NUTS-0).³⁵ Second, Table 11 states the declaration date of

³⁵ For the Netherlands, there are only national, i.e., no region-specific operational programmes and the official beneficiary lists do not contain information on their location. Partly, a NUTS-1 classification would be possible. However, we learn about the NUTS-2 region in which beneficiaries operate from ORBIS in two thirds of Dutch projects. To observations with an ORBIS match and a non-missing value for the postal code in ORBIS, we can assign a NUTS-3 region.

each OP, i.e., the reporting date of the list of beneficiaries or, if not provided by the managing authority, the date when we downloaded the list. Finally, Table 11 indicates the degree of detail of reported project sums: From only committed co-financing values (C) to EU and national public co-funding and private ineligible expenditure data (C_EU + C_NAT + I), and whether we know the value actually paid-out.

Robustness check

For a smaller sample of projects additional information from ORBIS is available. Table 12 presents the regression results including additional data, i.e., estimating the following:

$$\ln(TV_i) = \alpha + \psi R_i + \gamma F_i + \delta O_i + \rho T_i + \phi I_i + \kappa U_i + \lambda E_i + \mu Y_i + \varepsilon \quad (5)$$

where U_i is the sales volume of the beneficiary of project i , E_i the number of employees of the beneficiary (sales and number of employees corresponds to the last available observation in ORBIS) and Y_i the founding year of the beneficiary (in five brackets—before 1950, between 1950 and 1980, between 1980 and 2000, between 2000 and 2010 and after 2010). Furthermore, the regression contains all control variables included in Eq. (4), apart from the size of the beneficiary S_i as this ORBIS variable depends on the number of employees of the beneficiary E_i and its revenue U_i . In order to exploit the variance of relationships between project and beneficiary characteristics across regions, the regressions are not only done for the whole sample for which all of these variables are available. Further, we split the sample into (1) rural versus (2) urban regions as well as (3) less developed versus (4) developed (not classified as less developed by the 75% threshold applied to regional GDP relative to the EU average) regions. In those estimations, contrarily to previous ones, country-fixed effects are included instead of region-fixed effects.

Estimation results confirm that the size of the beneficiary matters for the value of a single project. Controlling for all other variables, higher revenues and more employees are associated with significantly higher single project values. When the number of employees of a beneficiary firm increases by 1%, the project it carries out is by 12% larger in terms of the project value. Taking into account whether a project and beneficiary, respectively, is located in an urban or rural region, this number amounts to 14% in rural and 17% in urban regions. Also in less developed regions this difference holds true. However, when considering only developed regions a one-percent-rise in employment is associated with only a slightly higher total project value (by 2%).

Furthermore, the age of the firm or institution receiving funds matters. Especially in less developed (as compared to other) and rural (as compared to urban) regions, younger beneficiaries which were founded after 2000 and after 2010 have projects with significantly higher total values than companies which were founded before that. Beneficiaries incorporated before 1980 carry out projects with the smallest total values.

Single projects' values in various industries (NACE Rev. 2 sectors) also differ between urban and rural as well as less developed and other regions. In some cases

Table 11 List of operational programmes (OP)

	OP Code	Fund	Obj.	NUTS	Committed	Paid-out	Declaration date
BG	2007BG051PO001	ESF	1	ORBIS	C	No	11.06.2015
BG	2007BG051PO002	ESF	1	–	–	–	–
BG	2007BG161PO001	ERDF	1	–	–	–	–
BG	2007BG161PO002	ERDF	1	–	–	–	–
BG	2007BG161PO003	ERDF	1	–	–	–	–
BG	2007BG161PO004	ERDF & CF	1	–	–	–	–
BG	2007BG161PO005	ERDF & CF	1	0	C	Yes	2.07.2013
BE	2007BE051PO001	ESF	1	ORBIS	C_EU + C_ NAT + I	Yes	10.2014
BE	2007BE052PO001	ESF	2	ORBIS	C_EU + C_ NAT + I	Yes	10.2014
BE	2007BE052PO002	ESF	2	ORBIS	C_EU + C_ NAT + I	Yes	10.2014
BE	2007BE052PO003	ESF	2	ORBIS	C_EU + C_ NAT + I	Yes	10.2014
BE	2007BE052PO004	ESF	2	ORBIS	C_EU + C_ NAT + I	Yes	10.2014
BE	2007BE052PO005	ESF	2	ORBIS	C_EU + C_ NAT + I	Yes	10.2014
BE	2007BE161PO001	ERDF	1	2	C_EU + C_ NAT + I	Yes	10.2014
BE	2007BE162PO001	ERDF	2	2	–	Yes	10.2014
BE	2007BE162PO002	ERDF	2	ORBIS	C	Yes	09.2014
BE	2007BE162PO003	ERDF	2	ORBIS	C_EU + C_ NAT + I	Yes	10.2014
CZ	2007CZ052PO001	ESF	2	2	C_EU	Yes	04.09.2015
CZ	2007CZ05UPO001	ESF	1 & 2	0	C_EU	Yes	04.09.2015
CZ	2007CZ05UPO002	ESF	1 & 2	0	C_EU	Yes	04.09.2015
CZ	2007CZ161PO001	ERDF	1	2	C_EU	Yes	04.09.2015
CZ	2007CZ161PO002	ERDF	1	2	C_EU	Yes	04.09.2015
CZ	2007CZ161PO004	ERDF	1	0	C_EU	Yes	04.09.2015
CZ	2007CZ161PO005	ERDF	1	2	C_EU	Yes	04.09.2015
CZ	2007CZ161PO006	ERDF & CF	1	0	C_EU	Yes	04.09.2015
CZ	2007CZ161PO007	ERDF & CF	1	0	C_EU	Yes	04.09.2015
CZ	2007CZ161PO008	ERDF	1	2	C_EU	Yes	04.09.2015
CZ	2007CZ161PO009	ERDF	1	2	C_EU	Yes	04.09.2015
CZ	2007CZ161PO010	ERDF	1	2	C_EU	Yes	04.09.2015
CZ	2007CZ161PO012	ERDF	1	0	C_EU	Yes	04.09.2015
CZ	2007CZ161PO013	ERDF	1	2	C_EU	Yes	04.09.2015
CZ	2007CZ162PO001	ERDF	2	2	C_EU	Yes	04.09.2015
CZ	2007CZ16UPO001	ERDF	1 & 2	0	C_EU	Yes	04.09.2015
CZ	2007CZ16UPO002	ERDF	1 & 2	0	C_EU	Yes	04.09.2015
DK	2007DK052PO001	ESF	2	2	–	Yes	08.09.2015
DK	2007DK162PO001	ERDF	2	2	–	Yes	08.09.2015

Table 11 (continued)

	OP Code	Fund	Obj.	NUTS	Committed	Paid-out	Declaration date
DE	2007DE051PO001	ESF	1	1	C (or paid-out)	Partly yes	31.12.2014
DE	2007DE051PO002	ESF	1	1	C (or paid-out)	Partly yes	31.12.2013
DE	2007DE051PO003	ESF	1	1	C (or paid-out)	Partly yes	31.03.2015
DE	2007DE051PO004	ESF	1	1	C (or paid-out)	Partly yes	06.2015
DE	2007DE051PO005	ESF	1	1	C	Yes	31.12.2014
DE	2007DE051PO006	ESF	1	1	C (or paid-out)	Partly yes	31.12.2014
DE	2007DE052PO001	ESF	2	1	C (or paid-out)	Partly yes	30.04.2014
DE	2007DE052PO002	ESF	2	1	C (or paid-out)	Partly yes	08.12.2014
DE	2007DE052PO003	ESF	2	1	C	Yes	31.12.2014
DE	2007DE052PO004	ESF	2	1	C (or paid-out)	Partly yes	06.11.2013
DE	2007DE052PO005	ESF	2	1	C (or paid-out)	Partly yes	31.12.2013
DE	2007DE052PO006	ESF	2	1	C	–	28.02.2014
DE	2007DE052PO007	ESF	2	1	C (or paid-out)	Partly yes	31.03.2015
DE	2007DE052PO008	ESF	2	1	C	–	31.12.2014
DE	2007DE052PO009	ESF	2	1	C	Partly yes	30.06.2014
DE	2007DE052PO010	ESF	2	1	C	–	05.10.2015
DE	2007DE052PO011	ESF	2	1	C (or paid-out)	Partly yes	06.2015
DE	2007DE05UPO001	ESF	1 & 2	0	C (or paid-out)	Yes	30.04.2015
DE	2007DE161PO001	ERDF	1	1	C (or paid-out)	Partly yes	31.12.2014
DE	2007DE161PO002	ERDF	1	1	C	Yes	31.12.2014
DE	2007DE161PO003	ERDF	1	1	C (or paid-out)	Partly yes	31.12.2013
DE	2007DE161PO004	ERDF	1	1	C (or paid-out)	Partly yes	06.2015
DE	2007DE161PO005	ERDF	1	0	C	–	31.12.2014
DE	2007DE161PO006	ERDF	1	1	C (or paid-out)	Partly yes	31.03.2015
DE	2007DE161PO007	ERDF	1	1	C	Yes	31.12.2014
DE	2007DE162PO001	ERDF	2	1	C (or paid-out)	Partly yes	31.12.2014
DE	2007DE162PO002	ERDF	2	1	C	Yes	31.12.2013
DE	2007DE162PO003	ERDF	2	1	C (or paid-out)	Partly yes	31.12.2014
DE	2007DE162PO004	ERDF	2	1	C	Yes	31.12.2014

Table 11 (continued)

	OP Code	Fund	Obj.	NUTS	Committed	Paid-out	Declaration date
DE	2007DE162PO005	ERDF	2	1	C (or paid-out)	Partly yes	16.04.2015
DE	2007DE162PO006	ERDF	2	1	C (or paid-out)	Partly yes	16.09.2015
DE	2007DE162PO007	ERDF	2	1	C	Yes	31.12.2014
DE	2007DE162PO008	ERDF	2	1	C (or paid-out)	Partly yes	31.12.2014
DE	2007DE162PO009	ERDF	2	1	C (or paid-out)	Partly yes	31.12.2014
DE	2007DE162PO010	ERDF	2	1	C (or paid-out)	Partly yes	31.03.2015
DE	2007DE162PO011	ERDF	2	1	C (or paid-out)	Partly yes	01.07.2015
EE	2007EE051PO001	ESF	1	2	C	Yes	31.08.2015
EE	2007EE161PO001	ERDF & CF	1	2	C	Yes	31.08.2015
EE	2007EE161PO002	ERDF & CF	1	2	C	Yes	31.08.2015
ES	2007ES051PO002	ESF	1	2	C	Yes	20.05.2013
ES	2007ES051PO003	ESF	1	2	C	Yes	20.05.2013
ES	2007ES051PO004	ESF	1	2	C	Yes	20.05.2013
ES	2007ES051PO005	ESF	1	2	C	Yes	25.07.2013
ES	2007ES051PO006	ESF	1	2	C	Yes	20.05.2013
ES	2007ES051PO007	ESF	1	2	C	Yes	20.05.2013
ES	2007ES051PO008	ESF	1	2	C	Yes	20.05.2013
ES	2007ES051PO009	ESF	1	2	C	Yes	20.05.2013
ES	2007ES052PO001	ESF	2	2	C	Yes	20.05.2013
ES	2007ES052PO002	ESF	2	2	C	Yes	20.05.2013
ES	2007ES052PO003	ESF	2	2	C	Yes	20.05.2013
ES	2007ES052PO004	ESF	2	2	C	Yes	20.05.2013
ES	2007ES052PO005	ESF	2	2	C	Yes	20.05.2013
ES	2007ES052PO006	ESF	2	2	C	Yes	20.05.2013
ES	2007ES052PO007	ESF	2	2	C	Yes	20.05.2013
ES	2007ES052PO008	ESF	2	2	C	Yes	20.05.2013
ES	2007ES052PO009	ESF	2	2	C	Yes	20.05.2013
ES	2007ES052PO010	ESF	2	2	C	Yes	20.05.2013
ES	2007ES052PO011	ESF	2	2	C	Yes	20.05.2013
ES	2007ES05UPO001	ESF	1 & 2	0	C	Yes	25.07.2013
ES	2007ES05UPO002	ESF	1 & 2	0	C	Yes	20.05.2013
ES	2007ES05UPO003	ESF	1 & 2	0	C	Yes	20.05.2013
ES	2007ES161PO001	ERDF	1	2	C	Yes	13.10.2015
ES	2007ES161PO002	ERDF	1	2	C	Yes	13.10.2015
ES	2007ES161PO003	ERDF	1	2	C	Yes	13.10.2015
ES	2007ES161PO004	ERDF	1	2	C	Yes	13.10.2015
ES	2007ES161PO005	ERDF	1	2	C	Yes	13.10.2015
ES	2007ES161PO006	ERDF	1	2	C	Yes	13.10.2015

Table 11 (continued)

	OP Code	Fund	Obj.	NUTS	Committed	Paid-out	Declaration date
ES	2007ES161PO007	ERDF	1	2	C	Yes	13.10.2015
ES	2007ES161PO008	ERDF	1	2	C	Yes	13.10.2015
ES	2007ES161PO009	ERDF & CF	1	0	C	Yes	13.10.2015
ES	2007ES162PO001	ERDF	2	2	C	Yes	01.02.2016
ES	2007ES162PO002	ERDF	2	2	C	Yes	13.10.2015
ES	2007ES162PO003	ERDF	2	2	C	Yes	13.10.2015
ES	2007ES162PO004	ERDF	2	2	C	Yes	13.10.2015
ES	2007ES162PO005	ERDF	2	2	C	Yes	13.10.2015
ES	2007ES162PO006	ERDF	2	2	C	Yes	13.10.2015
ES	2007ES162PO007	ERDF	2	2	C	Yes	13.10.2015
ES	2007ES162PO008	ERDF	2	2	C	Yes	13.10.2015
ES	2007ES162PO009	ERDF	2	2	C	Yes	13.10.2015
ES	2007ES162PO010	ERDF	2	2	C	Yes	13.10.2015
ES	2007ES162PO011	ERDF	2	2	C	Yes	13.10.2015
ES	2007ES16UPO001	ERDF	2	0	C	Yes	13.10.2015
ES	2007ES16UPO002	ERDF	1 & 2	0	C	Yes	31.10.2015
ES	2007ES16UPO003	ERDF	1 & 2	0	C	Yes	31.10.2015
FR	2007FR051PO001	ESF	1	2	C_EU + C_NAT	–	20.07.2016
FR	2007FR051PO002	ESF	1	2	C_EU + C_NAT	–	20.07.2016
FR	2007FR051PO003	ESF	1	2	C_EU + C_NAT	–	20.07.2016
FR	2007FR051PO004	ESF	1	2	C_EU + C_NAT	–	20.07.2016
FR	2007FR052PO001	ESF	2	2	C_EU + C_NAT	–	
FR	2007FR161PO001	ERDF	1	2	C_EU + C_NAT	–	20.07.2016
FR	2007FR161PO002	ERDF	1	2	C_EU + C_NAT	–	20.07.2016
FR	2007FR161PO003	ERDF	1	2	C_EU + C_NAT	–	20.07.2016
FR	2007FR161PO004	ERDF	1	2	C_EU + C_NAT	–	20.07.2016
FR	2007FR162PO005	ERDF	2	2	C_EU + C_NAT	–	20.07.2016
FR	2007FR162PO006	ERDF	2	2	C_EU + C_NAT	–	20.07.2016
FR	2007FR162PO007	ERDF	2	2	C_EU + C_NAT	–	20.07.2016
FR	2007FR162PO008	ERDF	2	2	C_EU + C_NAT	–	20.07.2016
FR	2007FR162PO009	ERDF	2	2	C_EU + C_NAT	–	20.07.2016

Table 11 (continued)

	OP Code	Fund	Obj.	NUTS	Committed	Paid-out	Declaration date
FR	2007FR162PO010	ERDF	2	2	C_EU + C_NAT	–	20.07.2016
FR	2007FR162PO011	ERDF	2	2	C_EU + C_NAT	–	20.07.2016
FR	2007FR162PO012	ERDF	2	2	C_EU + C_NAT	–	20.07.2016
FR	2007FR162PO013	ERDF	2	2	C_EU + C_NAT	–	20.07.2016
FR	2007FR162PO014	ERDF	2	2	C_EU + C_NAT	–	20.07.2016
FR	2007FR162PO015	ERDF	2	2	C_EU + C_NAT	–	20.07.2016
FR	2007FR162PO016	ERDF	2	2	C_EU + C_NAT	–	20.07.2016
FR	2007FR162PO017	ERDF	2	2	C_EU + C_NAT	–	20.07.2016
FR	2007FR162PO018	ERDF	2	2	C_EU + C_NAT	–	20.07.2016
FR	2007FR162PO019	ERDF	2	2	C_EU + C_NAT	–	20.07.2016
FR	2007FR162PO020	ERDF	2	2	C_EU + C_NAT	–	20.07.2016
FR	2007FR162PO021	ERDF	2	2	C_EU + C_NAT	–	20.07.2016
FR	2007FR162PO022	ERDF	2	2	C_EU + C_NAT	–	20.07.2016
HR	2007HR051PO001	ESF	1	ORBIS	C_EU + C_NAT	–	01.02.2015
HR	2007HR161PO001	CF	1	ORBIS	C_EU + C_NAT	–	01.07.2015
HR	2007HR161PO002	ERDF	1	ORBIS	C_EU + C_NAT	–	01.06.2015
HR	2007HR161PO003	ERDF	1	ORBIS	C_EU + C_NAT	–	10.06.2015
IE	2007IE052PO001	ESF	2	0	C or paid-out	–	19.09.2016
IE	2007IE162PO001	ERDF	2	2	C_EU + C_NAT	Yes	10.2015
IE	2007IE162PO002	ERDF	2	2	C	Yes (2)	09.2016
IT	2007IT051PO001	ESF	1	2	–	–	16.02.2015
IT	2007IT051PO002	ESF	1	2	C	Yes	22.10.2015
IT	2007IT051PO003	ESF	1	2	C	Yes	27.09.2012
IT	2007IT051PO004 (ITC2 & ITF5)	ESF	1	2	C	Yes	30.10.2014
IT	2007IT051PO005	ESF	1	2	C	–	30.12.2013
IT	2007IT051PO006	ESF	1	2	C	Yes	19.01.2015
IT	2007IT051PO007	ESF	1	–	–	–	–
IT	2007IT052PO001	ESF	2	2	C	Yes	31.12.2014

Table 11 (continued)

	OP Code	Fund	Obj.	NUTS	Committed	Paid-out	Declaration date
IT	2007IT052PO002	ESF	2	2	C	–	31.12.2014
IT	2007IT052PO003	ESF	2	2	C	Yes	31.12.2013
IT	2007IT052PO004	ESF	2	2	C	–	31.12.2014
IT	2007IT052PO005	ESF	2	2	C	Yes	31.12.2014
IT	2007IT052PO006	ESF	2	2	C	–	31.09.2015
IT	2007IT052PO007	ESF	2	2	C	Yes	31.12.2013
IT	2007IT052PO008	ESF	2	2	C	Yes	31.12.2014
IT	2007IT052PO009	ESF	2	2	C	Yes	19.05.2015
IT	2007IT052PO010	ESF	2	2	C	Yes	31.12.2014
IT	2007IT052PO011	ESF	2	2	C	Yes	05.10.2015
IT	2007IT052PO012	ESF	2	2	C	–	14.04.2015
IT	2007IT052PO013	ESF	2	2	C	–	16.01.2015
IT	2007IT052PO014	ESF	2	–	–	–	–
IT	2007IT052PO015	ESF	2	2	C_EU + C_ NAT + I	Yes	31.08.2015
IT	2007IT052PO016	ESF	2	2	C	Yes	31.12.2014
IT	2007IT052PO017	ESF	2	2	C	Yes	19.01.2015
IT	2007IT161PO001	ERDF	1	2	C	Yes	31.10.2015
IT	2007IT161PO002	ERDF	1	2	C	Yes	31.12.2014
IT	2007IT161PO003	ERDF	1	2	C_EU + C_NAT	Yes	31.08.2013
IT	2007IT161PO004	ERDF	1	–	–	–	–
IT	2007IT161PO005	ERDF	1	2	C	–	30.12.2015
IT	2007IT161PO006	ERDF	1	–	–	–	–
IT	2007IT161PO007	ERDF	1	2	C	Partly yes	14.10.2013
IT	2007IT161PO008	ERDF	1	2	C_EU + C_NAT	Yes	12.01.2016
IT	2007IT161PO009	ERDF	1	2	C_EU + C_NAT	Yes	11.11.2015
IT	2007IT161PO010	ERDF	1	2	C	Yes	31.05.2015
IT	2007IT161PO011	ERDF	1	2	C	Yes	28.02.2015
IT	2007IT161PO012	ERDF	1	2	C_EU + C_NAT	Yes (2)	31.12.2014
IT	2007IT162PO001	ERDF	2	2	C	Yes (2)	13.01.2016
IT	2007IT162PO002	ERDF	2	2	C	Partly yes	31.10.2014
IT	2007IT162PO003	ERDF	2	2	C_EU + C_NA	Yes (2)	30.09.2015
IT	2007IT162PO004	ERDF	2	2	C	Yes	22.06.2015
IT	2007IT162PO005	ERDF	2	2	C_EU + C_NAT	Yes	31.08.2015
IT	2007IT162PO006	ERDF	2	2	C	Yes	15.06.2015
IT	2007IT162PO007	ERDF	2	2	C_EU + C_NAT	Yes (2)	25.03.2014
IT	2007IT162PO008	ERDF	2	2	C_EU + C_NAT	Yes (2)	31.10.2013

Table 11 (continued)

	OP Code	Fund	Obj.	NUTS	Committed	Paid-out	Declaration date
IT	2007IT162PO009	ERDF	2	2	C	Yes	31.09.2015
IT	2007IT162PO010	ERDF	2	2	C	Yes	24.07.2015
IT	2007IT162PO011	ERDF	2	2	C	Yes	15.01.2016
IT	2007IT162PO012	ERDF	2	2	C	Yes	31.12.2014
IT	2007IT162PO013	ERDF	2	2	C	Partly yes	23.10.2015
IT	2007IT162PO014	ERDF	2	2	C	–	30.06.2013
IT	2007IT162PO015	ERDF	2	2	C_EU + C_NAT + I	Yes	30.06.2015
IT	2007IT162PO016	ERDF	2	2	C_EU + C_NAT	Yes (2)	20.11.2015
LV	2007LV051PO001	ESF	1	2	C_EU + C_NAT + I	Yes (2)	09.08.2015
LV	2007LV161PO001	ERDF	1	2	C_EU + C_NAT + I	Yes (2)	09.08.2015
LV	2007LV161PO002	ERDF & CF	1	2	C_EU + C_NAT + I	Yes (2)	09.08.2015
LT	2007LT051PO001	ESF	1	2	C_EU + C_NAT	Yes (2)	10.09.2015
LT	2007LT051PO002	ESF	1	2	C_EU + C_NAT	Yes (2)	10.09.2015
LT	2007LT161PO001	ERDF & CF	1	2	C_EU + C_NAT	Yes (2)	10.09.2015
LT	2007LT161PO002	ERDF & CF	1	2	C_EU + C_NAT	Yes (2)	10.09.2015
LU	2007LU052PO001	ESF	2	2	C_EU + C_NAT	–	12.2013
LU	2007LU162PO001	ERDF	2	2	C_EU + C_NAT	–	12.2014
MT	2007MT051PO001	ESF	1	2	C_EU + C_NAT	Yes (2)	09.11.2015
MT	2007MT161PO001	ERDF & CF	1	2	C_EU + C_NAT	Yes (2)	30.09.2015
NL	2007NL052PO001	ESF	2	0	C_EU + C_NAT + I	Yes	01.08.2016
NL ^a	2007NL162PO001	ERDF	2	0	C_EU + C_NAT + I	Yes	21.11.2016
NL	2007NL162PO002	ERDF	2	0	C_EU + C_NAT + I	Yes	21.11.2016
NL	2007NL162PO003	ERDF	2	0	C_EU + C_NAT + I	Yes	02.2015
NL	2007NL162PO004	ERDF	2	0	C_EU + C_NAT + I	Yes	21.11.2016
AT	2007AT051PO001	ESF	1	2	C	Yes	15.08.2015
AT	2007AT052PO001	ESF	2	0	–	Yes	20.11.2015
AT	2007AT161PO001	ERDF	1	2	C (or paid-out)	Partly yes	15.08.2015
AT	2007AT162PO001	ERDF	2	2	C	Yes	13.01.2015

Table 11 (continued)

	OP Code	Fund	Obj.	NUTS	Committed	Paid-out	Declaration date
AT	2007AT162PO002	ERDF	2	2	C	Yes	13.01.2015
AT	2007AT162PO003	ERDF	2	2	C	Yes	13.01.2015
AT	2007AT162PO004	ERDF	2	2	C	Yes	13.01.2015
AT	2007AT162PO005	ERDF	2	2	C	Yes	13.01.2015
AT	2007AT162PO006	ERDF	2	2	C	Yes	13.01.2015
AT	2007AT162PO007	ERDF	2	2	C	Yes	13.01.2015
AT	2007AT162PO008	ERDF	2	2	C	Yes	13.01.2015
PL	2007PL051PO001	ESF	1	2	C	Yes (2)	30.06.2015
PL	2007PL051PO001	ESF	1	2	C	Yes (2)	30.06.2015
PL	2007PL161PO001	ERDF	1	2	C	Yes (2)	30.06.2015
PL	2007PL161PO002	ERDF & CF	1	2	C	Yes (2)	30.06.2015
PL	2007PL161PO003	ERDF	1	2	C	Yes (2)	30.06.2015
PL	2007PL161PO004	ERDF	1	2	C	Yes (2)	30.06.2015
PL	2007PL161PO005	ERDF	1	2	C	Yes (2)	30.06.2015
PL	2007PL161PO006	ERDF	1	2	C	Yes (2)	30.06.2015
PL	2007PL161PO007	ERDF	1	2	C	Yes (2)	30.06.2015
PL	2007PL161PO008	ERDF	1	2	C	Yes (2)	30.06.2015
PL	2007PL161PO009	ERDF	1	2	C	Yes (2)	30.06.2015
PL	2007PL161PO010	ERDF	1	2	C	Yes (2)	30.06.2015
PL	2007PL161PO011	ERDF	1	2	C	Yes (2)	30.06.2015
PL	2007PL161PO012	ERDF	1	2	C	Yes (2)	30.06.2015
PL	2007PL161PO013	ERDF	1	2	C	Yes (2)	30.06.2015
PL	2007PL161PO014	ERDF	1	2	C	Yes (2)	30.06.2015
PL	2007PL161PO015	ERDF	1	2	C	Yes (2)	30.06.2015
PL	2007PL161PO016	ERDF	1	2	C	Yes (2)	30.06.2015
PL	2007PL161PO017	ERDF	1	2	C	Yes (2)	30.06.2015
PL	2007PL161PO018	ERDF	1	2	C	Yes (2)	30.06.2015
PL	2007PL161PO019	ERDF	1	2	C	Yes (2)	30.06.2015
PL	2007PL161PO020	ERDF	1	2	C	Yes (2)	30.06.2015
PT	2007PT051PO001	ESF	1	–	–	–	–
PT	2007PT052PO001	ESF	2	2	C_EU + C_NAT	–	02.11.2015
PT	2007PT05UPO001	ESF	1 & 2	0	C	–	28.01.2014
PT	2007PT05UPO001	ESF	1 & 2	0	C	–	28.01.2014
PT	2007PT05UPO002	ESF	1 & 2	0	1: C_EU + C_NAT, 2: C	–	12.2013
PT	2007PT161PO001	ERDF	1	0	C_EU + C_NAT	–	22.11.2015
PT	2007PT161PO002	ERDF	1	2	C_EU + C_NAT	–	31.07.2015
PT	2007PT161PO003	ERDF	1	2	C_EU + C_ NAT + I	–	31.07.2015
PT	2007PT161PO004	ERDF	1	2	C_EU + C_NAT	–	30.06.2015

Table 11 (continued)

	OP Code	Fund	Obj.	NUTS	Committed	Paid-out	Declaration date
PT	2007PT161PO005	ERDF	1	2	C_EU + C_NAT	–	30.06.2015
PT	2007PT161PO006	ERDF	1	2	C_EU + C_NAT	–	31.07.2015
PT	2007PT162PO001	ERDF	2	2	C_EU + C_NAT	–	31.07.2015
PT	2007PT162PO002	ERDF	2	2	C_EU + C_NAT + I	–	31.07.2015
PT	2007PT16UPO001	ERDF & CF	1	0	C_EU + C_NAT	–	30.09.2015
PT	2007PT16UPO002	ERDF	1 & 2	0	C_EU + C_NAT	–	30.11.2014
RO	2007RO051PO001	ESF	1	0	C_EU + C_NAT + I	–	31.07.2015
RO	2007RO051PO002	ESF	1	0	C_EU + C_NAT + I	–	31.07.2015
RO	2007RO161PO001	ERDF	1	0	C_EU + C_NAT + I	–	31.07.2015
RO	2007RO161PO002	ERDF	1	0	C_EU + C_NAT + I	–	31.07.2015
RO	2007RO161PO003	ERDF & CF	1	0	C_EU + C_NAT + I	–	31.07.2015
RO	2007RO161PO004	ERDF & CF	1	0	C_EU + C_NAT + I	–	31.07.2015
RO	2007RO161PO005	ERDF	1	0	C_EU + C_NAT + I	–	31.07.2015
SI	2007SI051PO001	ESF	1	2	C_EU + C_NAT	–	14.10.2015
SI	2007SI161PO001	ERDF	1	2	C_EU + C_NAT	–	14.10.2015
SI	2007SI161PO002	ERDF & CF	1	2	C_EU + C_NAT	–	14.10.2015
SK	2007SK05UPO001	ESF	1 & 2	2	C_EU + C_NAT + I	Yes (2)	30.06.2015
SK	2007SK05UPO002	ESF	1 & 2	2	C_EU + C_NAT + I	Yes (2)	30.06.2015
SK	2007SK161PO001	ERDF	1	2	C_EU + C_NAT + I	Yes (2)	30.06.2015
SK	2007SK161PO002	ERDF & CF	1	2	C_EU + C_NAT + I	Yes (2)	30.06.2015
SK	2007SK161PO003	ERDF	1	2	C_EU + C_NAT + I	Yes (2)	30.06.2015
SK	2007SK161PO004	ERDF & CF	1	2	C_EU + C_NAT + I	Yes (2)	30.06.2015
SK	2007SK161PO005	ERDF	1	2	C_EU + C_NAT + I	Yes (2)	30.06.2015
SK	2007SK161PO006	ERDF	1	2	C_EU + C_NAT + I	Yes (2)	30.06.2015

Table 11 (continued)

	OP Code	Fund	Obj.	NUTS	Committed	Paid-out	Declaration date
SK	2007SK161PO007	ERDF	1	2	C_EU + C_ NAT + I	Yes (2)	30.06.2015
SK	2007SK162PO001	ERDF	2	2	C_EU + C_ NAT + I	Yes (2)	30.06.2015
SK	2007SK16UPO001	ERDF	1 & 2	2	C_EU + C_ NAT + I	Yes (2)	30.06.2015
FI	2007FI052PO001	ESF	2	2	C_EU + C_NAT	Yes (2)	12.10.2015
FI	2007FI052PO002	ESF	2	2	C_EU + C_NAT	Yes (2)	12.10.2015
FI	2007FI162PO001	ERDF	2	2	C_EU + C_NAT	Yes (2)	12.10.2015
FI	2007FI162PO002	ERDF	2	2	C_EU + C_NAT	Yes (2)	12.10.2015
FI	2007FI162PO003	ERDF	2	2	C_EU + C_NAT	Yes (2)	12.10.2015
FI	2007FI162PO004	ERDF	2	2	C_EU + C_NAT	Yes (2)	12.10.2015
FI	2007FI162PO005	ERDF	2	–	–	–	–
SE	2007SE052PO001	ESF	2	2	C or paid-out	–	18.11.2015
SE ^b	2007SE162PO001	ERDF	2	0	C or paid-out	–	06.11.2015
SE	2007SE162PO002	ERDF	2	0	C or paid-out	–	06.11.2015
SE	2007SE162PO003	ERDF	2	0	C or paid-out	–	06.11.2015
SE	2007SE162PO004	ERDF	2	0	C or paid-out	–	06.11.2015
SE	2007SE162PO005	ERDF	2	0	C or paid-out	–	06.11.2015
SE	2007SE162PO006	ERDF	2	0	C or paid-out	–	06.11.2015
SE	2007SE162PO007	ERDF	2	0	C or paid-out	–	06.11.2015
SE	2007SE162PO008	ERDF	2	0	C or paid-out	–	06.11.2015
UK	2007UK051PO001	ESF	1	1	C_EU + C_NAT	–	30.04.2014
UK	2007UK051PO002	ESF	1	1	C_EU + C_NAT	–	25.08.2016
UK	2007UK052PO001	ESF	2	1	C_EU + C_NAT	–	25.08.2016
UK	2007UK052PO002	ESF	2	1	C_EU + C_NAT	–	30.04.2014
UK	2007UK052PO003	ESF	2	1	C_EU + C_ NAT + I	Yes	23.11.2016
UK	2007UK05UPO001	ESF	1 & 2	–	–	–	–
UK	2007UK161PO001	ERDF	1	1	C_EU + C_NAT	–	30.04.2014
UK	2007UK161PO002	ERDF	1	1	C_EU + C_NAT	–	25.08.2016
UK	2007UK161PO003	ERDF	1	1	C_EU + C_NAT	–	31.01.2015
UK	2007UK162PO001	ERDF	2	1	C_EU + C_NAT	–	30.04.2014

Table 11 (continued)

	OP Code	Fund	Obj.	NUTS	Committed	Paid-out	Declaration date
UK	2007UK162PO002	ERDF	2	1	C	–	01.05.2014
UK	2007UK162PO003	ERDF	2	1	C_EU + C_NAT + I	Partly yes (3)	08.10.2015
UK	2007UK162PO004	ERDF	2	1	C_EU + C_NAT + I	Yes (2)	01.01.2014
UK	2007UK162PO005	ERDF	2	1	C_EU + C_NAT	–	31.12.2013
UK	2007UK162PO006	ERDF	2	1	C_EU + C_NAT + I	Yes	01.10.2015
UK	2007UK162PO007	ERDF	2	1	C_EU + C_NAT	–	29.06.2016
UK	2007UK162PO008	ERDF	2	1	C	–	29.06.2016
UK	2007UK162PO009	ERDF	2	1	C_EU + C_NAT	–	29.06.2016
UK	2007UK162PO010	ERDF	2	1	C_EU + C_NAT	–	29.06.2016
UK	2007UK162PO011	ERDF	2	1	C_EU + C_NAT	–	31.01.2015
UK	2007UK162PO012	ERDF	2	1	C_EU + C_NAT	–	25.08.2016
UK	2007UK162PO012	ERDF	2	–	–	–	–

This table lists the operational programmes and information found in the official lists of beneficiaries provided by regional managing authorities or other regional or national authorities. The underlying list of operational programmes including their names can be downloaded at http://ec.europa.eu/regional_policy/en/policy/evaluations/data-for-research/ in the section “EU Budget commitments by fund by year and by programme” when selection the programming period 2007–2013. ^aProjects of 2007NL162PO002 and 2007NL162PO004 are part of the list of beneficiary for OP 2007NL162PO001. ^bUnequivocal assignment of ERDF projects to one of ERDF OPs is not possible. *C_EU* stands for committed EU co-financing, *C_NAT* for committed national co-financing, and *I* signifies ineligible cost, i.e., the cost carried by the beneficiary. Hence, *C_EU + C_NAT + I* means that the structure of the total project value is known. *C* means the a committed value, without declared partition across national and EU budget, is reported. A “Yes” in column *Paid-out* means that we have information on paid-out values. “Yes (2)” denotes that we know the partition of the paid-out amount into EU and national public co-financing. If the fifth and consequent columns do not contain any information, the dataset does not cover them as we have not found beneficiaries lists provided by the respective authorities. When the NUTS dimension given is “ORBIS”, it means that we have NUTS-2 information for those beneficiaries of the respective OP we could find and match with the ORBIS database. Objective 1 refers to the *Convergence* objective, Objective 2 to *Regional competitiveness and employment*

the difference might be due to location decisions of the companies and the regional sectoral structure. For example, the projects in *Financial and insurance activities* or *Other service activities* are larger in urban and developed regions. Moreover, projects with beneficiaries in *Water supply*, *Education* and *Energy* industries are larger in rural as compared to urban and (for the *Energy* industry only slightly smaller) in less developed as compared to higher-income (developed) regions.

Table 12 Regression coefficients with respect to project and all beneficiary characteristics

	General (Eq. 5)	Rural (Eq. 5)	Urban (Eq. 5)	Less dev. (Eq. 5)	Developed (Eq. 5)
No. of employees (log)	0.12***	0.14***	0.17***	0.15***	0.02***
Revenues (log)	0.08***	0.10***	0.03***	0.10***	0.11***
Foundation before 2010	+ 54%	+ 31%	− 0%	+ 94%	+ 31%
Foundation after 2010	+ 49%	+ 36%	+ 4%	+ 106%	+ 20%
Foundation before 2000	+ 49%	+ 19%	+ 13%	+ 86%	+ 25%
Foundation before 1980	+ 7%	+ 18%	− 28%	+ 85%	− 29%
Foundation before 1950	0%	0%	0%	0%	0%
Convergence	+ 23%	+ 44%	+ 6%	− 21%	+ 4%
CF	+ 123%	+ 149%	+ 127%	+ 64%	+ 89%
ERDF	+ 105%	+ 115%	+ 131%	+ 17%	+ 128%
ERDF or CF	+ 100%	+ 121%	+ 18%	+ 80%	+ 371%
ESF	0%	0%	0%	0%	0%
D—Energy (without water)	+ 116%	+ 121%	+ 114%	+ 92%	+ 94%
E—Water supply; sewerage, waste	+ 69%	+ 88%	+ 60%	+ 79%	+ 54%
U—Extraterr. organ. and bodies	+ 58%	+ 63%		+ 15%	+ 95%
L—Real estate activities	+ 49%	+ 51%	+ 42%	+ 16%	+ 46%
B—Mining and quarrying	+ 26%	+ 35%	+ 46%	− 0%	+ 40%
Q—Human health and social work	+ 20%	+ 36%	+ 14%	− 11%	+ 31%
S—Other service activities	+ 17%	+ 16%	+ 41%	− 17%	+ 30%
M—Prof., scientific and techn. act.	+ 3%	+ 9%	+ 1%	− 31%	+ 38%
R—Arts, entertainment and recr.	0%	0%	0%	0%	0%
P—Education	− 6%	+ 56%	− 1%	+ 41%	+ 10%
J—Information and communication	− 7%	+ 3%	− 5%	− 33%	+ 29%
K—Financial and insurance act.	− 10%	+ 24%	+ 52%	− 90%	+ 81%
H—Transportation and storage	− 13%	− 11%	− 4%	− 41%	+ 21%
I—Accommodation and food serv.	− 13%	+ 7%	− 30%	− 19%	− 3%
C—Manufacturing	− 19%	− 7%	− 17%	− 46%	+ 15%
N—Admin. and support serv.	− 20%	− 3%	− 19%	− 73%	+ 17%
F—Construction	− 39%	− 33%	− 18%	− 53%	− 29%
O—Public adm. and defense.	− 40%	− 38%	− 67%	− 56%	− 38%
G—Wholesale/retail; repair vehicles	− 47%	− 39%	− 34%	− 74%	− 23%
A—Agriculture, forestry and fishing	− 60%	− 56%	− 52%	− 90%	− 20%
T—Households as employers	− 104%	− 110%	− 46%	− 139%	− 99%
Other transport	+ 245%	+ 261%	+ 296%	+ 206%	+ 286%
Road	+ 202%	+ 269%	+ 372%	+ 194%	+ 55%
Rail	+ 160%	+ 83%	+ 245%	+ 98%	+ 173%
Culture heritage and tourism	+ 143%	+ 195%	+ 170%	+ 110%	+ 154%
Social inclusion	+ 55%	+ 110%	+ 58%	+ 14%	+ 87%
Innovation & RTD	+ 52%	+ 99%	+ 47%	+ 20%	+ 77%
Social infrastructure	+ 52%	+ 112%	+ 15%	+ 24%	− 38%
Environment	+ 44%	+ 72%	+ 76%	− 20%	+ 110%
Urban and territorial dimension	+ 22%	+ 92%	+ 112%	+ 24%	+ 87%

Table 12 (continued)

	General (Eq. 5)	Rural (Eq. 5)	Urban (Eq. 5)	Less dev. (Eq. 5)	Developed (Eq. 5)
IT services and infrastructure	0%	0%	0%	0%	0%
Labour market	- 1%	+ 100%	- 20%	- 69%	+ 33%
Other SME and business support	- 1%	+ 55%	- 18%	- 4%	+ 21%
Human capital	- 8%	+ 20%	- 25%	- 63%	+ 51%
Capacity building	- 43%	- 11%	- 120%	- 241%	+ 106%
Energy	- 79%	- 109%	+ 24%	- 189%	+ 78%
Number of obs.	174,624	78,570	50,223	70,116	77,445
Adjusted R ²	0.516	0.593	0.522	0.509	0.433

Coefficients of regressions based on Eq. (5). Dependent variable: logarithm of the total project value. Country-fixed effects are included. White-robust standard errors. Coefficients of respective dummy variable relative to the median category

When comparing the influence of theme on project values it becomes apparent that project size varies considerably between urban and rural (NUTS-3) regions for some of the themes. Projects corresponding to the *Rail*, *Other Transport*, *Environment*, *Urban and Territorial Dimension* and *Energy* themes are larger in urban (as compared to rural) and developed (compared to less developed) regions. Also *Road* projects are larger in urban regions which may be partly due to higher construction costs for infrastructure projects in urban areas. All other themes, especially those targeted at *Social Inclusion*, *Social Infrastructure*, *Labour Market* and *Other SME and Business Support*, are large in terms of conditional project values in rural regions. With regard to the development status of the NUTS-2 regions, results suggest that relative to *IT Services and Infrastructure* projects, just *Road* and *Social Infrastructure* are considerably larger in less developed than in higher-income regions.

Finally, the plot of the residual variation, which is not explained by project and firm characteristics, by country shows a very similar picture as in the baseline regression. The graph is provided in the Online-Appendix of this paper.

References

- Bachtler J, Mendez C (2007) Who governs EU cohesion policy? Deconstructing the reforms of the structural funds. *J Common Mark Stud* 45(3):535–564
- Bachtler J, Mendez C, Oraže H (2014) From conditionality to Europeanization in central and eastern Europe: administrative performance and capacity in cohesion policy. *Eur Plan Stud* 22(4):735–757
- Bachtrögler J (2016) On the effectiveness of EU structural funds during the great recession: estimates from a heterogeneous local average treatment effects framework. Working paper series, 230. Department of Economics, WU Vienna University of Economics and Business, Vienna
- Bachtrögler J, Fratesi U, Perucca G (2018) The influence of the local context on the implementation and impact of EU cohesion policy. *Reg Stud* (forthcoming)
- Barone G, David F, de Blasio G (2016) Boulevard of broken dreams: the ends of the EU funding (1997: Abruzzi, Italy). *Reg Sci Urban Econ* 60:31–38

- Becker SO, Egger PH, Von Ehrlich M (2012) Too much of a good thing? On the growth effects of the EU's regional policy. *Eur Econ Rev* 56(4):648–668
- Becker SO, Egger PH, von Ehrlich M (2013) Absorptive capacity and the growth and investment effects of regional transfers: a regression discontinuity design with heterogeneous treatment effects. *Am Econ J Econ Policy* 5(4):29–77
- Becker SO, Egger PH, von Ehrlich M (2018) Effects of EU regional policy: 1989–2013. *Reg Sci Urban Econ* 69:143–152
- Bernini C, Pellegrini G (2011) How are growth and productivity in private firms affected by public subsidy? Evidence from a regional policy. *Reg Sci Urban Econ* 41(3):253–265
- Bouvet F, Dall'Erba S (2010) European regional structural funds: how large is the influence of politics on the allocation process? *J Common Mark Stud* 48(3):501–528
- Breidenbach P, Mitze T, Schmidt CM (2016) EU structural funds and regional income convergence: a sobering experience. CEPR discussion paper no. 11210
- Cappelen A, Castellacci F, Fagerberg J, Verspagen B (2003) The impact of EU regional support on growth and convergence in the European Union. *J Common Mark Stud* 41(4):621–644
- Charron N, Dijkstra L, Lapuente V (2015) Mapping the regional divide in Europe: a measure for assessing quality of government in 206 European regions. *Soc Indic Res* 122(2):315–346
- Crisuolo C, Martin R, Overman H, Van Reenen J (2012) The causal effects of an industrial policy. NBER working paper no. 17842
- Dall'Erba S, Le Gallo J (2007) The impact of EU regional support on growth and employment. *Czech J Econ Financ* 57(7–8):325–340
- Dall'Erba S, Le Gallo J (2008) Regional convergence and the impact of European structural funds over 1989–1999: a spatial econometric analysis. *Papers Reg Sci* 87(2):219–244
- Dall'Erba S, Fang F (2017) Meta-analysis of the impact of European Union Structural Funds on regional growth. *Reg Stud* 51(6):822–832
- De Zwaan M, Merlevede B (2013) Regional policy and firm productivity. *ETSG* 2013—no. 377
- Dellmuth LM (2011) The cash divide: the allocation of European Union regional grants. *J Eur Public Policy* 18(7):1016–1033
- Dellmuth LM, Stoffel MF (2012) Distributive politics and intergovernmental transfers: the local allocation of European Union structural funds. *Eur Union Polit* 13(3):413–433
- Dijkstra L, Poelman H (2011) Regional typologies: a compilation. European Union, Paris
- European Commission (2006) Commission Regulation (EC) No. 1828/2006 of 8 December 2006, setting out rules for the implementation of Council Regulation (EC) No. 1083/2006 laying down general provisions on the European Regional Development Fund, the European Social Fund and the Cohesion Fund and of Regulation (EC) No. 1080/2006 of the European Parliament and of the Council on the European Regional Development Fund. *Official Journal of the European Union*
- European Commission (2011) Commission Regulation (EC) No. 31/2011 of 17 January 2011, amending annexes to Regulation (EC) No. 1059/2003 of the European Parliament and of the Council on the establishment of a common classification of territorial units for statistics (NUTS). *Official Journal of the European Union*
- European Commission (2017) European structural and investment funds. Retrieved on 03 May 2017 from http://ec.europa.eu/regional_policy/en/funding
- European Council (2006) Council Decision of 6 October 2006 on Community strategic guidelines on cohesion (2006/702/EC). European Council
- European Council (2006) Council Regulation (EC) No. 1083/2006 of 11 July 2006, laying down general provisions on the European Regional Development Fund, the European Social Fund and the Cohesion Fund and repealing Regulation (EC) No. 1260/1999. European Council
- Feinerer I, Hornik K, Meyer D (2008) Text mining infrastructure in R. *J Stat Softw* 25(1):1–54
- Ferrara AR, McCann P, Pellegrini G, Stelder D, Terribile F (2016) Assessing the impacts of cohesion policy on EU regions: a non-parametric analysis on interventions promoting research and innovation and transport accessibility. *Papers Reg Sci* 96(4):817–841
- Gagliardi L, Percoco M (2017) The impact of European Cohesion Policy in urban and rural regions. *Reg Stud* 51(6):857–868
- Hagen T, Mohl P (2009) Econometric evaluation of EU cohesion policy: a survey. ZEW discussion paper no. 09-052
- Hartsenko J, Sauga A (2012) Does financial support from the EU structural funds has an impact on the firms' performance: evidence from Estonia. In: *Proceedings of 30th international conference mathematical methods in economics*

- Horridge M, Rokicki B (2018) The impact of European Union accession on regional income convergence within the Visegrad countries. *Reg Stud* 53(4):503–515
- Joulin A, Grave E, Bojanowski P, Mikolov T (2016) Bag of tricks for efficient text classification. arXiv preprint [arXiv:1607.01759](https://arxiv.org/abs/1607.01759)
- Kalemlı-Ozcan S, Sorensen B, Villegas-Sanchez C, Volosovych V, Yesiltas S (2015) How to construct nationally representative firm level data from the ORBIS global database. NBER working paper no. 21558
- Kyriacou AP, Roca-Sagalés O (2012) The impact of EU structural funds on regional disparities within member states. *Environ Plan C Gov Policy* 30(27):267–281
- Le Gallo J, Dall’Erba S, Guillaın R (2011) The local versus global dilemma of the effects of structural funds. *Growth Change* 42(4):466–490
- Locatelli G, Mariani G, Sainati T, Greco M (2017) Corruption in public projects and megaprojects: there is an elephant in the room!. *Int J Project Manag* 35(3):252–268
- Maynou L, Saez M, Kyriacou A, Bacaria J (2016) The impact of structural and cohesion funds on Eurozone convergence, 1990–2010. *Reg Stud* 50(7):1127–1139
- Pellegrini G, Terribile F, Tarola O, Muccigrosso T, Busillo F (2013) Measuring the effects of European regional policy on economic growth: a regression discontinuity approach. *Papers Reg Sci* 92(1):217–233
- Percoco M (2013) Strategies of regional development in European regions: are they efficient? *Camb J Reg Econ Soc* 6:303–318
- Pienkowski J, Berkowitz P (2016) Econometric assessments of cohesion policy growth effects. In: Bachlter J, Berkowitz P, Hardy S, Muravska T (eds) *EU cohesion policy (open access): reassessing performance and direction: chapter 4*. Taylor & Francis, New York, pp 55–68
- Puga D (2002) European regional policies in light of recent location theories. *J Econ Geogr* 2(4):373–406
- R Core Team (2016) R: a language and environment for statistical computing. <http://www.R-project.org>. Accessed 3 June 2016
- Rodríguez-Pose A (2013) Do institutions matter for regional development? *Reg Stud* 47(7):1034–1047
- Rodríguez-Pose A, Fratesi U (2004) Between development and social policies: the impact of European structural funds in Objective 1 regions. *Reg Stud* 38(1):97–113
- Rodríguez-Pose A, Garcilazo E (2015) Quality of government and the returns of investment: examining the impact of cohesion expenditure in European regions. *Reg Stud* 49(8):1274–1290
- Stone M (1974) Cross-validatory choice and assessment of statistical predictions. *J Roy Stat Soc Ser B* 36:111–147
- Tosun J (2014) Absorption of regional funds: a comparative analysis. *JCMS J Common Mark Stud* 52(2):371–387
- Ward JH Jr (1963) Hierarchical grouping to optimize an objective function. *J Am Stat Assoc* 58(301):236–244

Affiliations

Julia Bachtrögl^{1,2} · Christoph Hammer² · Wolf Heinrich Reuter³ · Florian Schwendinger²

Christoph Hammer
christoph.hammer@wu.ac.at

Wolf Heinrich Reuter
wolf.reuter@svr-wirtschaft.de

Florian Schwendinger
florian.schwendinger@wu.ac.at

¹ Austrian Institute of Economic Research, Arsenal, Object 20, 1030 Vienna, Austria

² Vienna University of Economics and Business, Welthandelsplatz 1, 1020 Vienna, Austria

³ German Council of Economic Experts, Gustav-Stresemann-Ring 11, 65189 Wiesbaden, Germany