Acquisition of European research funds and its effect on international scientific collaboration

Jarno Hoekman[1]*, Thomas Scherngell[2], Koen Frenken[3], Robert Tijssen[4]

[1] *j.hoekman@tue.nl*

School of Innovation Sciences; Eindhoven University of Technology (The Netherlands)

Urban and Regional research centre Utrecht, Utrecht University (The Netherlands)

Tel: +31 040 247 4638

Fax: +31 040 244 4602

*corresponding author

[2] *thomas.scherngell@ait.ac.at*

AIT Austrian Institute of Technology (Austria)

[3] *k.frenken@tue.nl*

School of Innovation Sciences; Eindhoven University of Technology (The Netherlands)

[4] *tijssen@cwts.leidenuniv.nl*

Centre for Science and Technology Studies (CWTS), Leiden University, Leiden (The Netherlands)

Abstract: The Framework Programmes (FPs) funded by the European Commission support transnational research collaborations in order to make the European Research Area more competitive. Some have raised concerns, however, that the FPs compromise the cohesion policies aimed at reducing income disparities between European regions. Our research objective is *first*, to investigate whether existing scientific collaborations between EU regions – as captured by co-publications – are conducive for acquiring FP funding, and *second*, to study the effects of funding on subsequent co-publication activity between regions. Our results indicate that previous co-publications only have a minor effect on being funded. We also find that the effect of funding on co-publication activity is especially significant for regional pairs that did not intensively co-publish before participation. The results suggest that the returns to FP funding are highest when involving scientifically lagging regions. In this respect, the current FP policy is in line with cohesion policy.

Keywords: Research collaboration; Framework Programmes; European Union; NUTS2; Lisbon Agenda; gravity equation

JEL: O38; R12

Manuscript word count (excl. references): 7,513 words

1. Introduction

Despite the pervasive trends towards the globalisation of knowledge production, research policy is still driven mainly by national budgets and objectives. One of the major exceptions is the European Union, where at the Lisbon Council in 2000, the Heads of State signed up to develop a European Research Area (ERA). The objective of ERA policy is to improve the competitiveness and coordination of research (funding) activities at regional, national and EU levels. The Framework Programmes (FPs) of the European Commission (EC) constitute the main instrument to achieve this goal. They are specifically designed to pool resources and promote international R&D collaboration between the EU member states by enabling and intensifying interactions among researchers. The final goal is to stimulate knowledge creation and diffusion as prime sources for sustainable economic growth in the long run (Foray 2004; Romer 1990).

From its inception there has been much concern that the policy objectives pursued as part of the Lisbon strategy would compromise the cohesion objective of the European Union (Sharp 1998; Hoekman et al. 2009; Begg 2010). After all, Europe's research policies are not intended to intervene in the European scientific and technological landscape at large, but to bundle resources with the purpose of supporting collaborative efforts between 'excellent' actors in a few strategic scientific fields. Following common perceptions about the instrumentality of research collaboration (Katz and Martin 1997; Sonnenwald 2004), this strategy is expected to induce economies of scale, to avoid duplication of research efforts, and to enhance the competitiveness of the European territory as a whole vis-à-vis its main global competitors. Given the current unequal distribution of scientific and technological capabilities across the European landscape it can be reasonably inferred that such strategies

may well disproportionally support high-performance core regions, possibly even at the expense of peripheral regions. It is well known that scientific and technological activities show strong natural tendencies to spatially concentrate (Audretsch and Feldman 1996; Moreno et al. 2005). The creation and diffusion of knowledge supported by networks also follows geographical logics by its nature, favoring short distance interactions over longer distance ones (Adams et al. 2005; Fischer et al. 2006, Breschi and Lissoni 2009; Hoekman et al. 2009). Both tendencies might hamper the possibilities for peripheral actors to connect at a distance. In Europe these forces are arguably particularly strong, as distinct national and regional systems persist and countries maintain their own strategies next to the European-wide Lisbon Agenda (Crescenzi et al. 2007).

Recently, economic geographers have started to pay attention to the importance of multiple 'scales of innovation' (Bunnell and Coe 2001) including 'global pipelines' that provide regions with access to distant knowledge sources (Bathelt et al. 2004). Following this terminology, the FPs can be understood as a deliberate sttempt to create a European scale of innovation by financially supporting collaboration structures that can function as global pipelines. Yet, we currently lack an understanding of the effectiveness of such policy strategies changing the dominant spatial scales of knowledge production.

This study intends to depict a promising direction to fill this research gap by relating participation in the FP funding scheme to co-publication output as a generally accepted proxy for collaborative knowledge production (Frenken et al. 2009). The objective is first, to investigate whether existing international scientific collaborations between EU regions – as captured by co-publications in scientific journals – are conducive for the acquisition of FP funding, and second, to measure the effects of FP funding on subsequent co-publication

activity between EU regions. Should peripheral actors indeed have difficulties in connecting to more central ones at a distance, we will observe that acquisitions and effects of the FPs are disproportionally concentrated in already existing networks of scientific activity.

We address these research objectives by exploiting a unique time-series in which FP participation and FP network structures – as captured by joint participation of organisations in FP projects – are linked to publication output and co-publication networks at the macro level of 254 regions within 25 European countries. We develop a thematic concordance between scientific fields and FP thematic areas, which is needed to establish the link between FP networks and co-publication networks in different thematic fields. Using this concordance, we are able to analyse for three broad thematic areas to what extent existing co-publication networks predict subsequent FP funding acquisition and to what extent FP funds, in turn, affect subsequent co-publication patterns.

The remainder of the paper is organised as follows. First, we elaborate on our theoretical assumptions and pay explicit attention to the role of regions in Europe's research policies. We subsequently introduce the data in section 3 and the concordance between scientific fields and FP thematic areas in Section 4. Section 5 spells out the empirical framework. In section 6 we present our results. Finally, in section 7 we discuss the implications of our findings, limitations of our study and potential directions for further research.

2. EU Framework Programmes and the regions

The emergence of a systematic Research and Technological Development (RTD) policy at the European level can be traced back to the 1980s when the first multi-annual Framework

Programme was implemented. As the name suggests it is conceived as a common *framework* under which EU RTD policies should be organised and as a *programme* that lasts several years to make possible long term investments in specific strategic areas such as ICTs, sustainable development, biotechnology and energy. From its inception, the FPs have provided funds for transnational networks of researchers in order to overcome impediments to international research collaboration.

Traditionally, public support for RTD has been justified by making reference to market failures stemming from the inherent uncertainty of research and the public nature of knowledge (Arrow 1959; Foray 2004). During the early years of the FPs, this basic rationale was complemented by interactive and non-linear conceptualizations of the research system (Kline and Rosenberg 1986; Lundvall 1988) emphasizing that public interventions should not just provide additional resources but rather focus on solving 'system failures', including fostering collaborative learning, strengthening linkages between dissimilar actors and facilitating the diffusion of knowledge (Breschi and Malerba 2009). Progressively, attention has been given to specificities of innovation systems, be them regional (Cooke et al. 1997), sectoral (Malerba 2002) or technological (Hekkert et al. 2007).

More recently, the objectives of European RTD policies have become centred around '*a new strategic goal: to become the most competitive and dynamic knowledge-based economy in the world'* (European Council 2000, I.A.5), which has become publicly known as the 'Lisbon Agenda'. As a major pillar to achieve this goal, the mobilizing concept of the European Research Area (ERA) has been embraced. Using the scientific and technological performance of the United States and Japan as a benchmark, the European Commission postulates that *'the situation concerning research is worrying. Without concerted action this could lead to a loss*

of growth and competiveness in an increasingly globalised economy' (Commission 2000, p. 4). To improve this situation it is specifically stated that '*research activities at national and Union level must be better integrated in order to make them as efficient and innovative as possible, and to ensure that Europe offers attractive prospects to its best brains'* (European Council 2000, I.A.12). The objectives set out in the Commissions' communication 'Towards a European Research Area' provided the starting point for such an ongoing endeavour (Commission 2000) and the objectives were re-emphasized in more recent communications, including the launch of Europe's 2020 agenda (Commission 2007a; Commission 2010).[1]

The justification for ERA policy is that world-class '*centres of excellence*' (Commission 2007a, p. 14) already exist in Europe across a wide range of research fields. Yet, these centres often remain loosely connected and their expertise is not always sufficiently known and accesible across Europe. The integration of these '*centres of excellence*' in long-term self-organised R&D consortia should therefore facilitate the creation of a critical mass of research performing entities that possess a collaborative attitude and converge on strategic goals (Breschi and Malerba 2009). Given the expected instrumentality of research collaboration, this process in turn will induce economies of scale and scope, avoid duplication of research efforts and enhance the competitiveness of the European territory as a whole (Katz and Martin 1997; Sonnenwald 2004).

[1] It was not the first time that the Commission diagnosed weaknesses in Europe's research and industrial base vis-à-vis their main competitors (i.e. United States and Japan). Both the 1993 White Paper on Growth, Competitiveness and Employment (Commission 1993) and the 1995 Green Paper on Innovation (Commission 1995) already summarized a series of weaknesses under three distinct headings. *'The first weakness is financial. The community invests proportionally less than its competitors in research and technological development (...). A second weakness is the lack of coordination at various levels of the research and technological development activities, programmes and strategies in Europe (...). The greatest weakness however, is the comparatively limited capacity to convert scientific breakthroughs and technological achievements into industrial and commercial successes'* (Commission 1993, chapter 4). The reality of these weaknesses and their development over time remain the focus of academic debate. See on Europe's underinvestment in R&D (Pavitt 2000, Duchene et al. 2009); on fragmentation (Luukonen and Medeva 2009; Hoekman et al. 2010); on the 'European paradox' (Dosi et al. 2006).

To some extent a European core of research performing entities is already visible in the research system, in which leading city-regions that are located in close geographical vicinity to one other (e.g. London, Paris, Randstad Region, Milan, Munich, Geneva) are disproportionally connected (Frenken et al. 2007; Hoekman et al. 2009, Matthiessen et al. 2010). In the case of scientific activity, Figure 1 and Figure 2 which are based on the data described in the next section, support this observation. Scientific publication output per capita (Figure 1) is concentrated in a group of 'core' regions located in a Western European axis stretching south-east from London towards Rome, in Scandinavian regions and in some large city-regions located in other parts of Europe (e.g. Berlin, Budapest, Glasgow/Edinburgh, Madrid, Vienna). This confirms the well known observation that scientific activities show strong natural tendencies to spatially concentrate (Audretsch and Feldman 1996; Moreno et al. 2005). Scientific collaboration networks as proxied by co-publication counts (Figure 2) also show that only a small group of regional hubs take central positions in the European scientific collaboration networks (e.g. Berlin, London, Paris, the Randstad Region). These hub-regions inter-connect the dense national networks, in particular, those in France, Germany, Italy, the Netherlands, Spain and the UK. The localized nature of the networks reflect the spatial barriers to collaborative knowledge production, as captured by distance decays and national border effects in previous studies (Hoekman et al. 2009, 2010). It has been argued that these barriers are even more pronounced in the European context, which is characterised by historically fragmented national systems and regional identities (Crescenzi et al. 2007). Up to now, national governments continue to invest in strengthening such local or national scientific and technological capacities, and only partially align their goals and interests with the European research and innovation policies, including the Lisbon Agenda (Banchoff 2002; Guzetti 2009).

The existence of a core-periphery structure in scientific and technological knowledge production has raised concerns regarding possible conflicts between the objectives of the Lisbon strategy and the objectives of cohesion policy, which has also been noted by the European Commission at an early stage (Commission 2001). Cohesion policy aims to reduce income disparities between Europe's poorest regions and the rest of Europe. Hence, the two objectives can be incompatible, insofar as the establishment of ERA is likely to generate disproportionate benefits for the existing core of research performing entities.

Indeed, some studies that have analyzed the topological structures of the FP networks stress the complexities of combining cohesion and competitiveness objectives by showing that it is difficult for unconnected actors to acquire a central position in FP funding networks. Breschi and Cusmano (2004) and Autant-Bernard et al. (2007) analyze the social network structures among FP participants and find that the funded collaboration networks are dominated by a small *'oligarchic core'* of research actors, whose central network positions in the program have only strengthened over the successive funding rounds. Many participants in current FP projects were thus already participating in previous FPs (Paier and Scherngell 2010), and it can be easily inferred from this that participants from new member states will have difficulties to enter the FP network. Scherngell and Barber (2009, 2010) have studied the importance of spatial and other impediments for collaboration within the fifth FP and conclude that – despite the intended European character – the number of links between organisations in the fifth FP funding scheme tends to decay with geographical distance and language barriers. Similar findings have been reported by Maggioni and Uberti (2009) and Balland (2010). The evidence suggests that unconnected actors at a distance exhibit difficulties in acquiring funding within the FPs.

It is important to consider however, that these studies treat the FPs in isolation, without taking into account their complex interactions with the wider scientific, technological and economic contexts. An analysis of the early FPs shows regarding this issue that the FP funding favored core regions only in absolute terms. Peripheral countries still managed to acquire more funding relative to their total R&D capacities in line with the cohesion objective of the European Commission at the time (Sharp 1998; Clarysse and Muldur 2001). We take this issue a step further in doing a relational analysis, studying whether the number of previous co-publications between two regions is predictive of acquisition of FP funding by these two regions. Given the trade-off between competitiveness and cohesion in EU policy we are especially interested whether a disproportionate amount of funding flows to regional pairs within the existing European core of research performing entities.

An answer to our first research objective is also instrumental in dealing with our second research objective, which is to measure the effects of FP funding on establishing new international co-publication activities. Considering the expected effect of FP projects we can safely assume that the volume of FP funding will have at least some positive effect on co-publication activity because the FP funding scheme provides 'behavioural additionality' in this respect (Luukonen 2000). The additionality of FP funding is apparent from the fact that co-publications are much more likely to occur within the same country. In the European context these barriers tend to be quite pervasive and have only partially broken down over the last decade (Hoekman et al. 2010). Collaboration networks that require to be organised in

international large-scale R&D consortia are therefore very unlikely to emerge in a similar structure without strategic interventions.[2]

Hence, the question here is not whether FP funding has an effect on co-publication activity *per se* but rather whether regions differ in their ability to transform FP funding into international co-publication output. Research performing entities in core regions generally have better regional and national funding opportunities (Banchoff 2002; Dosi et al. 2006). Given the high levels of local funding for core regions, additional FP funding may generate only few additional co-publications. That is, FP funding may well substitute for these local funding sources in core regions, while they provide additionality for peripheral regions. However, if, indeed, researchers in core regions are, on average, more productive in publishing than their colleagues in peripheral areas for whatever reason, one would observe the effect of FP funding on co-publications to be larger for core regions than for peripheral regions.

3. Data

To test how acquisitions and effects of FP projectss are distributed over the European territory, we link data from two different sources: (i) research articles indexed by the *Web of Science* database, (ii) FP joint projects participations extracted from the *EUPRO database*.

[2] This observation is confirmed by the experience of participants who often mention the establishment of international partnerships as a main benefit of participation in the FPs (Luukonen 2000). Effects of FP funding on international scientific collaboration networks might also be substantial due to a number of other reasons. Among them are matching requirements by national governments or research organisations, which alone may double the volume of funding. It has also been documented that in some successful cases the FPs have played an agenda setting role, which may have induced multiplier effects on domestically funded research (Arnold et al. 2005). Moreover, a micro level study has shown the importance of FP funding in stimulating the productivity and collaborative behavior of researchers as captured by co-publications (Defazio et al. 2009). Given these observations, we might observe evolving international scientific collaboration networks that are closely associated with funding provided by the FPs.

The Web of Science database (WoS) is a bibliographical database produced by Thomson Reuters, indexing approximately 12,000 sources worldwide and considered to be one of the most comprehensive and reliable sources of information on basic research activity across all countries and fields of science. Its indexed research articles all occur in peer-reviewed journals. The journals are selected on the basis of a minimum quality assessment carried out by Thomson Reuters. We choose to analyse all research articles contained in the WoS, published in the period 2000-2007 and containing at least one European author-affiliate address. All publications are assigned to 22 disciplines as defined by *Netherlands* Observatory of Science and Technology (NOWT) and based on aggregations of the journal categories listed in WoS.[3]

The EUPRO database presently comprises data on all funded FP projects and participating organisations. It contains systematic information on project objectives and achievements, project costs, project funding and contract type as well as on the participating organisations including the full name, the full address and the type of the organisation (Roediger-Schluga and Barber 2006). Currently, data for the first to the fifth FP are complete, whereas sixth FP projects are covered up to 2006. For this analysis, we choose to extract all projects for three pairs of FP thematic areas (see section 4) that run in the fifth and sixth FP (referred to as FP5 and FP6 from now on). Because of the specific time-series on publications and the limited availability of FP6 data we focus on all projects that have been funded in the period 2000-2005.

[3] NOWT Science and Technology Indicators reports are available at www.nowt.nl. We exclude the 13 social science disciplines because the coverage of ISI is less in this domain and because their contribution to the FPs is rather small.

We restrict the analysis to a specific set of three related thematic areas in FP5 and FP6 (see section 4). The rationale for this focus is threefold: (i) the thematic areas receive the lion's share of funding within the FPs. More specifically, in FP5 they mutually make up 72.5% of funding, whereas in FP6 63.3% of funding is allocated within those themes (Expert Group 2009), (ii) one of the major goals of the thematic areas is their scientific and technological impact, implying that scientific publications can be considered a significant and meaningful output of the funded projects[4], and (iii) knowledge production within the thematic areas explicitly concentrates on collaborative actions which allows for a sound comparison not only with the activity of regions but also with their scientific collaboration networks.

As we adopt a spatial perspective we are interested in the locations involved in research production and collaboration. Accordingly, all institutional addresses on research articles and FP project description are uniquely assigned to European regions on the basis of city names and postal codes. More specifically, regions are defined by the hierarchical NUTS classification with each organisation being assigned to one out of 254 NUTS2 regions in 25 countries in Europe (Commission 2007b). The 25 countries include all countries of the European Union plus Norway and Switzerland, but excluding Romania, Bulgaria, Cyprus and Malta. In most countries these intra-national regions have administrative authority, although for five small countries they are defined at the national level.

International research collaboration is defined as a pair of different organisation addresses that occur in the same research article or on the same FP project description but are located in different countries. In our procedure, we count each international region-pair that occurs

[4] The number of publications may in the future also be used as an indicator for the evaluation of projects.

simultaneously in the by-line of an article or simultaneously on the project description.[5] Obviously, an article or project may contain multiple international region pairs depending on the number of different regional addresses that appear on a publication or project description. We subsequently aggregate the count of all international collaborative activities in the datasets into region-by-region matrices that denote the number of collaborations between region *i* and region *j* (*i*, *j* = 1, …, *n*) in year *t* located in different countries. As collaborations are undirected the final dataset contains 29,679 unique regional pairs for every year *t*.

4. Concordance

The funded projects within the FPs tend to focus on a few thematic areas. A systematic study of acquisition and effect, in which total FP participation is juxtaposed with total co-publication activities therefore runs the risk of serious estimation biases, as scientific collaboration networks may differ considerably across fields. In the following we therefore briefly report on the establishment of a concordance table that links the various thematic areas in the FPs to the scientific fields they target. An extensive description of the procedure and the dataset is provided in Appendix A.

As of 2009, the indexed publications in the various citation databases that are part of Web of Science can be searched on grant activity and funding acknowledgements (i.e. funding agencies and grant numbers). We use this tool to develop a query that searches in the funding acknowledgement texts for names and abbreviations related to European institutions, EC RTD policies and FPs. We retrieved the funding acknowledgement text and the journal source for every identified publication.

[5] The counts exclude multiple occurrences of similar regions within the same article or project description.

We subsequently build an algorithm that searches for unique grant numbers and call abbreviations of FP projects - available in the EUPRO database - in the acknowledgement texts of the publications. The algorithm searches for names of the various sub-programmes and for parts of 'model case' structures of grant numbers in FP5 and FP6, which - when mentioned appropriately - contain a unique programme identifier, a year and a contract number. In those cases where the algorithm links a research publication to a FP sub-programme we always checked manually whether the match was correct or false. Based on this procedure, 8,235 publication records could be assigned to a specific thematic area within FP5 or FP6. It turns out that in most thematic areas scientific publications are a frequently occurring output. Even the thematic areas of FP5 – running from 1998 to 2002 - still produced almost 1,400 publications in 2009.

As a third step, we aggregate for every thematic area the funded publications to one out of 22 scientific fields as defined by the CWTS/NOWT classification (Tijssen et al. 2010). Thus, a distribution of publication output over scientific fields is obtained for every FP thematic area. We then correlate these distributions to identify which thematic areas in FP5 and FP6 have a similar scientific profile. Following this procedure, we can reasonably infer three broad thematic areas (i.e. distinct sets of related programs in FP5 and FP6). (1) Scientific output of FP5-EESD and FP6-SUSTDEV (from now on SUSTDEV) is concentrated in 'Earth Sciences and Technology' and 'Environmental Science and Technology'; (2) Life science programs FP5-QOL and FP6-LSH, where scientific output over scientific fields is mainly concentrated in the scientific fields 'Basic Life Sciences' and 'Clinical Medicine' (from now on LIFESCIENCE); (3) FP5-IST and FP6-IST focus on information and communication

technologies and their output appears mainly in physical science and computer science outlets (from now on ICT).

As a last data-processing step, we adjust all regional publication and co-publication counts to create proxies of scientific output and scientific collaboration networks that resemble the scientific profile of the FP thematic areas. More specifically, for the three broad thematic areas (i.e. SUSTDEV, LIFESCIENCE and ICT) all 22 scientific fields are assigned weights. The weights equal for every broad thematic area the share of funded publications in a particular scientific field in the total number of funded publications for that broad thematic area multiplied by 22. We subsequently multiply for every scientific field, both regional publication and inter-regional co-publication counts with their respective weights and aggregate over 22 scientific fields. An overview of the broad thematic areas and their exact weights is provided in Table 1.

5. Research design

Model 1: Acquisition of FP funding

The first empirical objective of the paper is to investigate whether existing scientific collaboration networks – as captured by co-publications in scientific journals – are conducive for the acquisition of FP funding. To address this objective, we model inter-regional FP participation between region *i* and *j* by employing a cross-sectional spatial interaction model of the gravity type. These models have become a workhorse for the statistical analysis of aggregate compositions of human interactions ranging from trade, traffic flows and telephone calls, to marriages, museum trips and money flows (Sen and Smith 1995). In the context of research collaboration, the gravity model has often been used to explain the intensity of

research collaboration among European regions (Hoekman et al. 2009; 2010; Maggioni and Uberti 2009; Scherngell and Barber 2009; 2010). We adopt a cross-sectional Poisson regression design where the dependent variable is the number of international FP participations between region *i* and region *j*. This applies to each of the three broad thematic areas and is based on projects that started in the first three years of FP6 (2003-2005).

Following common practice in the application of gravity models, we first randomize the model by including weights – also referred to as origin and destination variables – for the total number of FP participation of region *i* and the total number of FP participation of region *j*. In case the acquisition of FP projects would be perfectly random between regions, only these weights would become significant with a value approximating one. Hence, all other significant effects should be interpreted as deviations from complete random acquisition of FP funding between European regions.

Besides controlling for origin and destination effects, the main component of spatial interaction models is the separation function that measures the separation between region *i* and *j*. It is naturally and common practice to assume that the geographical distance between *i* and *j* is an important factor influencing the probability of interactions. Thus, we follow this assumption and control for geographical distance effects.[6]

In the context of our first objective, the focus of interest is on the influence of existing scientific collaboration networks on joint acquisition of FP funding. We introduce the number of co-publications between region *i* and *j* in the three years prior to the start of FP6 (2000-

[6] In order to do so a measure is introduced that captures geographical distance between region *i* and region *j* in terms of the great circle distance.

2002) as an explanatory variable. In case a disproportional amount of funding flows to the existing core of research performing entities we expect a positive effect of these co-publications networks on the likelihood that region *i* and *j* acquire FP funding. In this case previous scientific collaborations thus aid in obtaining funding from the FPs, for instance because of existing collaboration routines and mutual trust.

We add some further control variables. First, we expect that being located in the same language area also increases the likelihood of participating in the FPs and therefore include a dummy variable denoting whether the same language is spoken in two regions (0) or not (1). *Second*, we estimate whether previous participation in the FPs increases the likelihood of participating in FP6 (Breschi and Cusmano 2004; Autant-Bernard et al. 2007). More specifically, we include for every regional pair the number of times they simultaneously occurred on a project description in a related FP5 thematic area, and the number of times a regional pair jointly participated in an unrelated FP5 thematic area.[7] An overview of all included independent variables is provided in Table 2.

Model 2: Effect of FP funding

The second empirical objective of this paper is to measure the effects of FP funding on the occurrence of international scientific collaboration networks. We explain the number of international co-publications between region *i* and region *j* after joint participation in FP

[7] For instance, when the number of links between region *i* and region *j* in FP6-SUSTDEV is the dependent variable, the number of joint project participation in FP5-EESD is considered *previous related participation*, whereas the aggregate number of links in FP5-IST and FP5-QOL between region *i* and *j* are considered as *previous unrelated links*. Based on this distinction, we try to assess whether the influence of previous participation is solely an outcome of a general international collaborative orientation (including expertise in European project participation), or a broad thematic area specific outcome due to proximities or similarities between FP partners.

projects, while controlling – amongst other factors – for the number of international co-publications before joint participation.

We exploit the panel nature of our dataset by regressing international co-publication counts between region *i* and region *j* in a given year (2003-2007) on a set of year-specific explanatory variables. As our observations are at the regional link level (between region *i* and region *j*), we again rely on a spatial interaction model of the gravity type (Sen and Smith 1995). This means that we introduce region-specific weights which are equal to the number of publications of region *i* and the number of publications of region *j* in the three years before we measure the effect on co-publication activity. As in the first model a Poisson regressive framework is employed, but this time a negative binomial variant is estimated due to unobserved heterogeneity between the (*i*, *j*)-region pairs leading to overdispersion (see Long and Freese 2001 for details).

The variable of main interest is the amount of funding that regional pairs receive, which is measured as the number of joint acquisitions of FP funding. If an effect of FP funding on the production of international co-publications is present we expect this variable to be positive and significant. We use a three-year-moving-window in which we include the cumulative amount of projects in which both region *i* and region *j* participated in the previous three years.

One can also expect that significant amounts of co-publication activity take place on the base of previously established connections (i.e. repeated ties). In our estimation framework we therefore include a measure of past performance that controls for the cumulative number of established co-publication relations between region *i* and *j* in the three prior years. In other

words, this is a measure of the existing scientific research collaboration networks between region *i* and region *j*.

As we are particularly interested whether the effects of funding are equally distributed over the European territory, we include an interaction term between the amount of funding and existing scientific collaboration networks as captured by past performance. By this, we are able to observe whether regional pairs that are better connected profit more or less from the acquisition of funding than regional pairs that are less well connected. All interacted variables are centred on their mean value, which implies that we observe whether funding has an effect when regional pairs have average numbers of existing scientific research collaborations.

Similar to the analysis on funding acquisition, we also introduce a set of *geographical controls* that may explain the rate and direction of international co-publication output among regions. Thus, we also include in this *model the geographical distance between* region *i* and region *j* and a dummy variable whether the same language is spoken (0) or not (1) in two regions.

After controlling for region-specific effects (i.e. number of publications) and some regional pair specific effect that are fixed over time (i.e. distance and language), we assume that all remaining time-invariant factors influencing international co-publication activity are random. Thus, we rely on a random effects variant of our gravity panel model to be estimated (see Fidrmuc 2009). An overview of all included independent variables is provided in Table 5.

6. Results

Model 1: Acquisition of funding

Table 3 reports descriptive statistics and correlations for the included variables. Both the number of existing co-publication relations and the acquisition of FP5 funding are highly correlated with the acquisition of FP6 funding. The acquisition of FP6 funding also decays significantly with distance and when a language border is crossed, but these decays are less strong than those observed for the scientific co-publication networks.

We present the cross-section Poisson regression explaining the number of joint projects in which both region *i* and region *j* participate in Table 4. Model 1A shows that in two of the three thematic areas (i.e. SUSTDEV and ICT) the number of existing co-publication relations exerts positive and significant influence on the amount of funding regional links receive, albeit with a very small coefficient. We also confirm that participation in FP6 tends to decay with geographical distance and that there is some evidence that FP participation decreases when a language border is crossed (cf. Scherngell and Barber 2009, 2010).

In Model 1B we include the effect of previous participation in FP5 on FP6 participation. Regional links that already participated in FP5 have a higher likelihood of participating in FP6 for all thematic areas. We also show that this is especially the case for participation in related programs, whereas experience in unrelated programs does not influence FP6 participation. Furthermore, we also note that in Model 2 the effect of existing co-publication activity on FP participation vanishes, with the possible exception of SUSTDEV which has a small and not very significant coefficient of 0.010 (p=0.093). The effect of existing co-publication activity is much smaller than the effect of previous participation in FP5 on participation in FP6, even when comparing the coefficients of existing co-publication activity

in Model 1A with the coefficient of previous FP participation in Model 1B.[8] The main outcome of the analysis thus suggests that within the FP funding scheme, experience in previous FPs is especially important for receiving FP funding, whereas previous co-publications only have a minor effect, if any, on being funded in FP6.

The results also suggest that better networked regions are not disproportionally selected into the funding scheme. This is important given that in our subsequent analysis of the effect of FP funding we need to make sure that potential effects of the FPs are indeed due to treatment (i.e. funding) and not just to the selection of better performing regions (see for instance Busom 2000). Since regional pairs that perform better in terms of co-publication output do not seem to receive disproportionally more funding than any other regional pair, we can start to analyze the effect of funding on subsequent co-publication activity indeed as a treatment effect.

Model 2: Effect of FP funding

The second analysis looks at the determinants of co-publication activity and, particularly, the effect of FP funding. As a first descriptive analysis, we looked at regions without any previous co-publications in the period 2000-2002. For these previously unconnected regions, Figure 3 shows the average number of co-publications after 2002 for different amounts of funding, for all three broad thematic areas. It becomes clear that funding has a positive effect, as Figure 3 shows that regional pairs with FP funding produce more joint publications than regional pairs without FP funding.

[8] Note that the coefficients are to be interpreted as elasticities.

The effect of FP funding on co-publication activity, however, should be assessed taking into account a number of control variables. In particular, we take into account the total publication activity in each region, their joint co-publication activity, the geographical distance between them, and language barriers. Descriptive statistics and correlation matrices of these variables are presented in Table 6.

In Model 2A and 2B we first explain the number of international co-publications between region *i* and *j*, for those regional pairs that did not have any co-publication output in the three previous years. For all three thematic areas, almost 50% of all regional pairs are part of this subset. Model 2A is a baseline model that contains only control variables. The controls exhibit significant effects on the number of international co-publications that are produced between regional pairs that did not previously co-publish. This means that the higher the number of publications in region *i* and in region *j*, the higher the likelihood that new international scientific research collaborations will follow. There are also spatial effects that explain the creation of new international scientific research collaborations. More precisely, the likelihood that international regional pairs that did not previously interact start co-publishing, decays with geographical distance, and decreases when a language border is crossed. Model 2B introduces the number of joint acquisitions of funding of region *i* and region *j* as an explanatory variable. Sign and coefficient of the amount of funding are positive and significant which confirms – at least for the group of regional pairs that did not previously co-publish – that there is a positive effect of participation in the FPs on the creation of new scientific research collaborations. This effect occurs independently of other included variables as their sign and significance are not affected by the inclusion of funding in the model.

From Model 2C onwards we present the analysis based on the entire sample of regional pairs, irrespective of whether the regional pair co-published in the past. This makes model 2C similar to model 2B with the exception that in model 2C we include all regional pairs in the analysis. We again find a significant positive effect of funding on the number of international co-publications, suggesting that funding aids in establishing new scientific collaborations for all regional pairs and not for only those that did not co-publish in the past. However, the coefficients of funding are significantly lower than in Model 2B when applying a 95% confidence interval.

In Model 2D we also control for the past performance of regional pairs by introducing the number of co-publications between region *i* and *j* in the previous 3 years as an explanatory variable. Judging from the coefficients, the number of co-publications have a strong and significant effect on the establishment of international co-publications between region *i* and *j*. Furthermore, the effect of funding on the establishment of international co-publications decreases even further with the inclusion of this control variable. Comparing the effect of funding for the sample of regional pairs that did not co-publish in the past (Model 2B) with the effect of funding given average levels of co-publication activity in the past (Model 2D), we note that the effect has now decreased to almost half of its initial value. Based on the outcome of Model 2C and Model 2D we therefore hypothesize that the effect of funding may well be moderated by the extent to which regions already co-published in the past.

Model 2E shows the results when the interaction term between previous research collaborations and amount of funding is included. Note that all variables are centered on the mean, which implies that we should interpret the coefficient of funding as the effect of funding given average levels of co-publications and *vice versa*. In Model 2E the coefficient

of funding takes on a similar value as in Model 2B, suggesting that funding indeed increases international co-publishing for regional pairs with average levels of scientific research collaboration networks. In addition, as in all previous models the number of prior international co-publications also exhibit positive and significant influence on international co-publishing, which effect is - in this case - conditional on average levels of funding.

Turning to the interaction variable we observe however that it has a negative and significant sign for all three broad thematic areas. This shows that the more a regional pair co-published before, the less the impact of FP funding on subsequent co-publication activity. That is, for regional pairs that already intensively co-published, FP funding does not provide much additionality. This suggests that for regional pairs that intensively co-published in the past, FP is not a driver for the establishment of new international co-publication activities but rather act as a substitute for other sources of funding.

We also carried out two robustness checks. First, we repeated the analysis on intervals of the data in which we distinguish between groups of regional pairs with approximately similar amounts of prior collaborations. In this case funding only becomes positively significant for samples of regional pairs that have low numbers of prior collaborations, whereas funding becomes insignificant or sometimes even negatively significant for samples with high numbers of prior collaborations. Second, we estimated models with moving windows of up to five years (i.e. funding and co-publication relations in the five years before time of measurement), because in some cases the production of a co-publication might take longer than three years after the start of a FP project. The results obtained from these analyses led us to the same conclusions.

7. Discussion

International research collaboration is believed to be beneficial for the production and diffusion of knowledge. Yet, cross-border collaboration is still significantly hampered by the dominance of national research systems. In the European context, the Framework Programmes have been designed to overcome this 'system failure'. They intend to create a *'better coordinated and integrated European Research Area'* by funding transnational collaboration networks between researchers, in order to make the European scientific and technological landscape more coherent and competitive. However, serious concerns have been voiced over the extent to which the available funds are disproportionally absorbed by established research actors which are typically located in Europe's core of scientific regions. If so, it can be argued that in the pursuit of competitiveness, cohesion objectives will easily be compromised, given the difficulties that peripheral actors face in connecting at a distance.

The analysis presented in this paper provides little empirical evidence to support this claim. Using an extensive regionalised dataset of scientific co-publication counts and FP participation counts, we studied whether acquisition and effect of FP funding are disproportionally concentrated in the already existing core of research performing entities. We find that the core regions that co-publish frequently do not receive a disproportionate amount of funding from the FPs. Our analysis also showed that the more a regional pair co-published before, the less the additional effect of FP funding on subsequent co-publication activity. This suggests that the effect of the FP funding scheme is subject to decreasing returns to scale and may substitute for other sources of funding. Put differently, the returns to FP funding seem to be the highest when involving lagging regions.

Although our findings underline the relevance of the FPs in supporting Europe's cohesion objective, the dynamics of research collaboration networks at large do not necessarily exhibit a clearly visible tendency towards a higher degree of cohesion. We showed that - next to the importance of FP funding - geographical factors (e.g. distance and language) remain significant barriers for establishing new international scientific collaboration networks. Scientific research collaboration networks are also more often established on the basis of already existing collaborations between actors. Given this topology of the European scientific landscape, the role of the FPs is especially effective in establishing ties between previously poorly connected regions, rather than in further strengthening existing ties between centers of excellence. Our finding thus sheds a critical light on the excellence rhetoric that has dominated the science policy discourse during the recent past. Given the current geography of science in Europe, the pursued *excellence* policies of the EU have a rather cohesive outcome.

References

Adams, J.D., Grant, G.C., Clemmons, J.R., Stephan, P.E. (2005) Scientific teams and institutional collaborations: evidence from US universities, 1981-1999. *Research Policy*, 34: 259-285.

Arnold, E., Clark, J., Muscio, A. (2005) What the evaluation record tells us about European Union Framework Programme performance. *Science and Public Policy*, 32: 385-397.

Arrow, K. (1959) Economic welfare and the allocation of resources for invention. *The Rand Corporation*, 15 december 1959.

Audretsch, D., Feldman, M. (1996) R&D spillovers and the geography of innovation and production. *American Economic Review*, 86: 630-640.

Autant-Bernard, Billand, P., Frachisse, P., Massard, N. (2007) Social distance versus spatial distance in R&D cooperation: empirical evidence from European collaboration choices in micro and nanotechnologies. *Papers in Regional Science*, 86: 495-519.

Balland, P.A. (2009) Proximity and the Evolution of Collaboration Networks: Evidence from R&D Projects within the GNSS Industry. *Regional Studies*, forthcoming.

Banchoff, T. (2002) Institutions, inertia and European Union research policy. *Journal of Common Market Studies*, 40: 1-21.

Bathelt, H., Malmberg, A., Maskell, P. (2004) Clusters and knowledge: local buzz, global pipelines and the process of knowledge creation. *Progress in Human Geography,* 28: 31-56.

Begg, I. (2010) Cohesion or confusion: a policy searching for objectives. *Journal of European Integration*, 32: 77-96.

Breschi, S., Cusmano, L. (2004) Unveiling the Texture of a European Research Area: Emergence of Oligarchic Networks under EU Framework Programmes. *International Journal of Technology Management*, 27: 747-772.

Breschi, S., Lissoni, F. (2009) Mobility of skilled workers and co-invention networks: an anataomy of local;izaed knowledge flows. *Journal of Economic Geography*, 9: 439-468.

Breschi, S., Malerba, F. (2009) ERA and the role of networks. In: deLanghe, H., Muldur, U., Soete, L. (eds.), *European Science and Technology Policy: Towards Integration or Fragmentation?* Edward Elgar: Cheltenham, pp.

Bunnell, T., Coe, N.M. (2001) Spaces and scales of innovation. *Progress in Human Geography*, 25: 569-589.

Busom, I. (2000) An empirical evaluation of the effects of R&D subsidies. *Economics of Innovation and New Technology,* 9: 111-148.

Clarysse, B., Muldur U. (2001) Regional cohesion in Europe? An analysis how EU public RTD support influences the techno-economic regional landscape. *Research Policy*, 30: 275-296.

Commission of the European Communities (1993) *White paper on growth, competitiveness and employment.* COM (93), 700 final, Brussels, 5 December 1993

Commission of the European Communities (1995) *Green paper on innovation.* COM(95), 688 final, Brussels, 20 December 1995.

Commission of the European Communities (2000) *Towards a European Research Area*. COM(2000) 6 final, Brussels, 18 January 2000.

Commission of the European Communities (2001) *The Regional Dimension of the European Research Area*. COM(2001) 549 final, Brussels, 3 October 2001.

Commission of the European Communities (2007a) *Green paper 'The European Research Area: New Perspectives'*. {SEC(2007) 412}, *COM(2007)161* final, Brussels, 4 April 2007.

Commission of the European Communities (2007b) Commission regulations (EC) no 105/2007. *Official Journal of the European Union*, I.39, 10 February 2007.

Commission of the European Communities (2010) *Europe 2020: a strategy for smart, sustainable and inclusive growth*. *COM(2010)2020* final, Brussels, 3 March 2010.

Cooke, P., Gomez-Uranga, M., Etxebarria, G. (1997) Regional innovation systems: institutional and organizational dimensions. *Research Policy*, 26(4-5): 475-491.

Crescenzi, R., Rodriguez-Pose, A., Storper, M. (2007) The territorial dynamics of innovation: a Europe-United States comprative analysis. *Journal of Economic Geography*, 7: 673-709.

Defazio, D., Lockett, A., Wright, M. (2009) Funding incentives, collaborative dynamics and scientific productivity: Evidence from the EU Framework Program. *Research Policy*, 38: 293-305.

Dosi, G., Llerena, P., Labini, M.S. (2006) The relationships between science, technologies and their industrial exploitation: an illustration through the myths and realities of the so-called 'European Paradox'. *Research Policy*, 35: 1450-1464.

Duchene, V., Lykogianni, E., Verbeek, A. (2009) The EU R&D under-investment: patters in R&D expenditure and financing. In: deLanghe, H., Muldur, U., Soete, L. (eds.), *European Science and Technology Policy: Towards Integration or Fragmentation?* Edward Elgar: Cheltenham, pp .

European Council (2000) *Presidency conclusions*. Lisbon European Council, 23-24 March 2000.

Expert group (2009) *Evaluation of the Sixth Framework Programme for research and technological development 2002-2006*. Report of the Expert Group.

Fidrmuc, J. (2009) Gravity models in integrated panels. *Empirical Economics* (37): 435-446.

Fischer M.M., Scherngell T. and Jansenberger E.M. (2006) The geography of knowledge spillovers between high-technology firms in Europe. Evidence from a spatial interaction modelling perspective. *Geographical Analysis* 38 (3): 288-309

Foray, D. (2004) *The Economics of Knowledge*. MIT Press: London.

Frenken, K., Hardeman, S, Hoekman J. (2009) Spatial Scientometrics: towards a cumulative research program. *Journal of Informetrics*, 3: 222-232.

Guzetti, L. (2009) The European Research Area idea in the history of community policy-making. In: deLanghe, H., Muldur, U., Soete, L. (eds.), *European science and technology policy: towards integration or fragmentation?* Edward Elgar: Cheltenham.

Hekkert, M.P., Suurs, R.A.A., Negro, S.O., Kuhlmann, S., Smits, R.E.H.M. (2007) Functions of innovation systems: A new approach for analysing technological change. *Technological Forecasting and Social Change*, 74: 413-432.

Hoekman, J., Frenken, K., Van Oort, F. (2009) The geography of collaborative knowledge production in Europe. *The Annals of Regional Science*, 43: 721-738.

Hoekman, J., Frenken, K., Tijssen, R. (2010) Research collaboration at a distance: changing spatial patterns of research collaboration in Europe. *Research Policy*, 39: 662-673.

Katz, J.S., Martin, B.R. (1997) What is research collaboration? *Research Policy*, 26: 1-18.

Kline, S.J., Rosenberg, N. (1986) An overview of innovation. In: R. Landau, N. Rosenberg (eds.), *The Positive Sum Strategy: Harnessing Technology for Economic Growth*. Washington DC: National Academy Press, pp..

Long, J.S., Freese, J. (2001) *Regression models for categorical dependent variables using STATA*. College Station, Texas: Stata Press.

Lundvall, BA (1988) Innovation as an intetractive process: from user-produce interaction to the national system of innovation. In: Dosi, G., Freeman, C., Nelson, R.R., Silverberg G. and Soete, L. (eds.), *Technical change and Economic Theory*, London: Pinter Publications, pp.

Luukonen, T. (2000) Additionality of EU Framework Programmes. *Research Policy*, 6: 711-724.

Luukonen, T., Medeva, N. (2009) Towards understanding integration in research and research policy. *Research Policy*, 39: 674-684.

Maggioni, M., Uberti, T. (2009) Knowledge networks across Europe: which distance matters? *The Annals of Regional Science* 43: 691-720.

Malerba, F. (2002) Sectoral systems of innovation and production. *Research Policy*, 31: 247-264.

Matthiessen, C.W., Schwarz, A.W., Find, S. (2010) World cities of scientific knowledge: systems, networks and potential dynamics. An analysis based on bibliometric indicators. *Urban Studies*, 47: 1879-1897.

Moreno, R., Paci, R., Usai, S. (2005) Geographical and sector clusters of innovation in Europe. *The Annals of Regional Science*, 39: 715-739.

Paier, M., Scherngell, T. (2010) Determinants of collaboration in European R&D networks: Empirical evidence from a discrete choice model. *Industry and Innovation*, forthcoming.

Pavitt, K. (2000) Why European Union funding of academic research should be increased: a radical proposal. *Research Policy*, 27: 455-460.

Roediger-Schluga, T., Barber, M. (2006) The structure of R&D collaborations networks in the European Framework Programmes: data prcessing, network construction and selected results. *International Journal of Foresight Innovation Policy*, 4: 321-347

Romer, P.M (1990) Endogenous technological change. *Journal of Political Economy,* 98: 71–102.

Scherngell, T. Barber, M. (2009) Spatial interaction modelling of cross-region R&D collaborations: empirical evidence from the 5th EU framework programme. *Papers in Regional Science*, 88: 531-546.

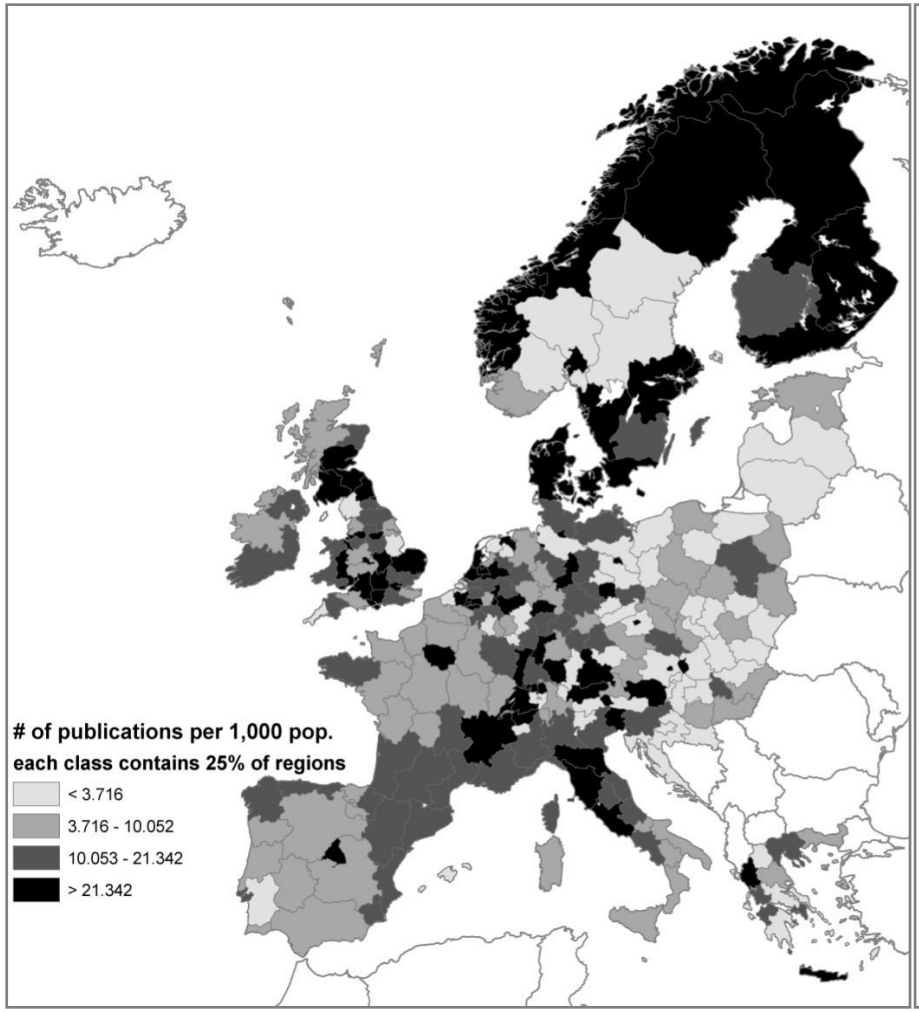
Scherngell, T., Barber, M. (2010) Distinct spatial characteristics of industrial and public research collaborations: Evidence from the 5th EU Framework Programme. *The Annals of Regional Science*, forthcoming.

Sharp, M. (1998) Competitiveness and cohesion. Are the two compatible? *Research Policy*, 27: 569-588.

Sen, A., Smith, T.E. (1995) *Gravity Models of Spatial Interaction Behavior*. New York: Springer.

Sonnenwald, D. (2004) Research collaboration. *Annual Review of Information Science and Technology*, 41: 643-681.

Tijssen, R.J.W., Hollanders, H., Kanerva, M., van Leeuwen, T.N., Nederhof, A.J., Van den Berg, P. (2010) Wetenschaps en Technologie-indicatoren 2010, Netherlands Observatorium van Wetenschap en Technologie (NOWT). *Report for the Netherlands Ministry of Education, Science and Culture*.



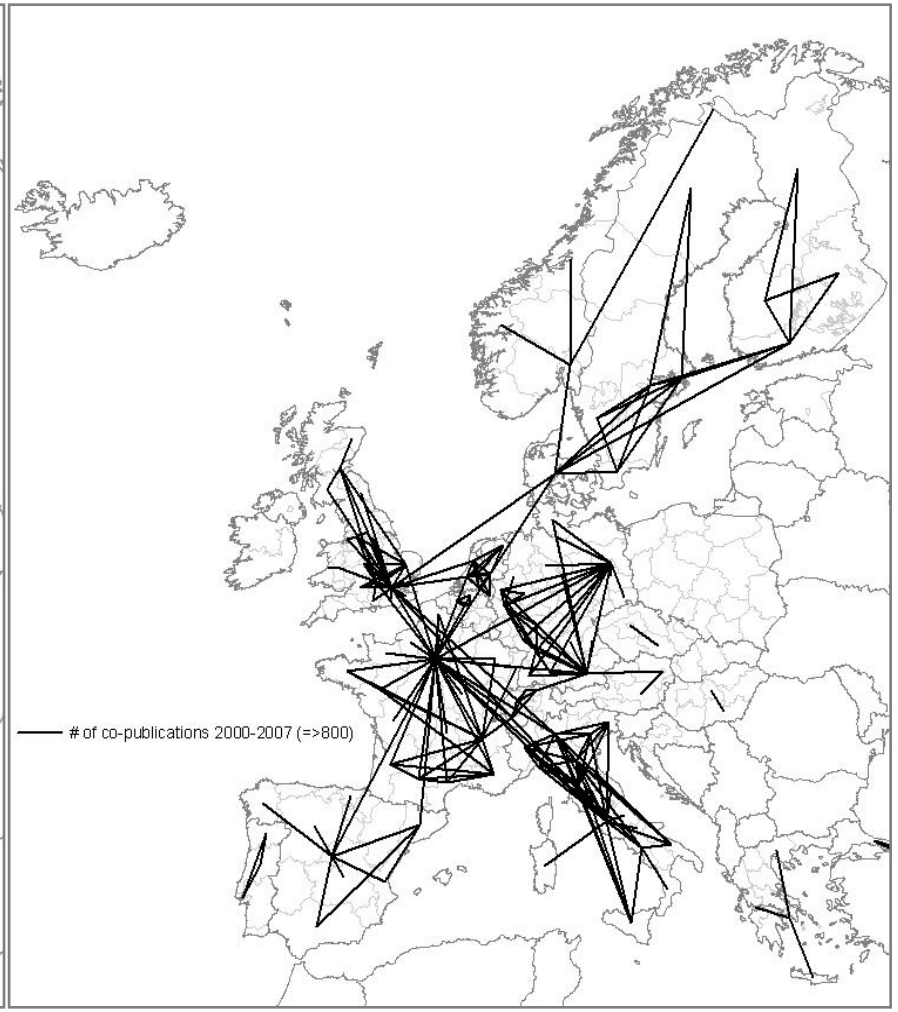
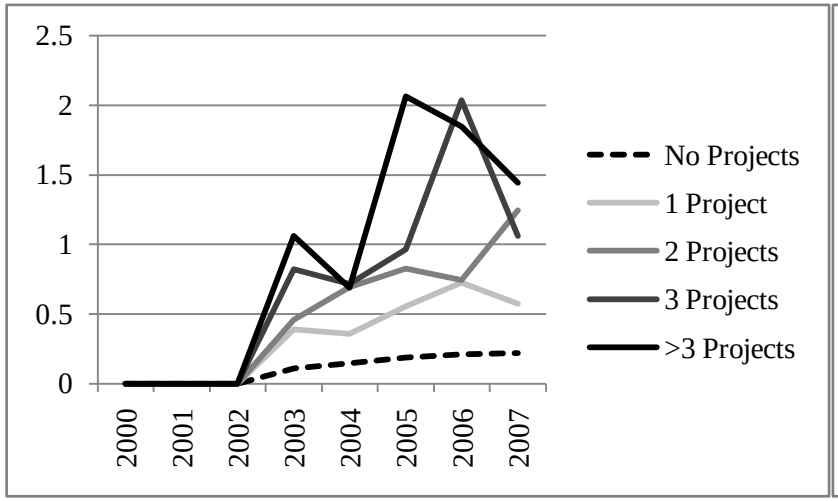
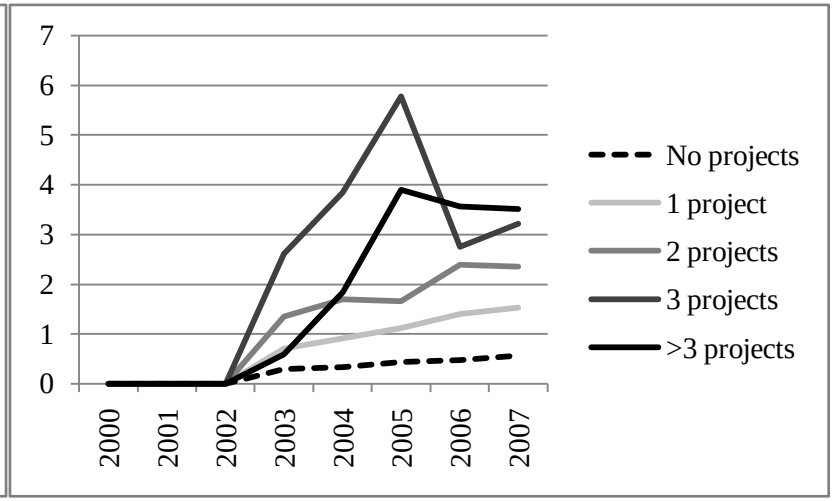


Figure 1 (left) and Figure 2 (right): Total number of publications per capita and co-publications in the period 2000-2007

Figure 3: Average number of co-publications (y-axis) for regional pairs that did not co-publish before joint participation in the FPs. SUSTDEV (upper-left), LIFESCIENCE (upper-right) and ICT (lower-left).





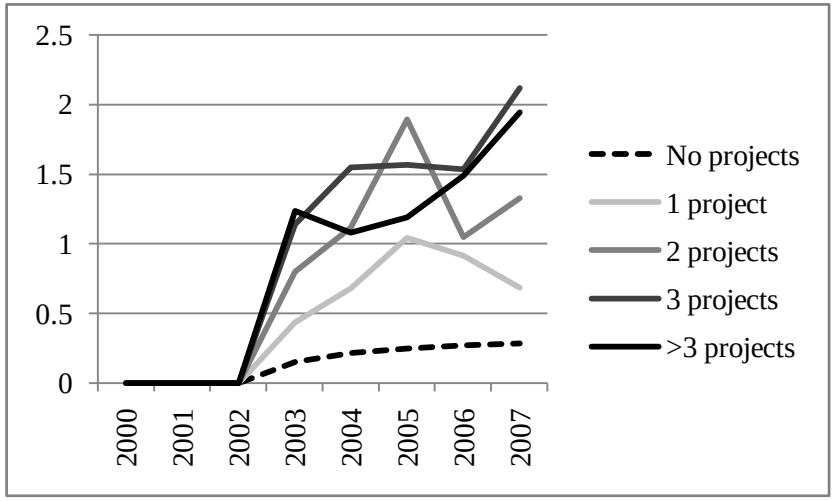


Table 1: Weights used for computing publication and co-publication counts*

Broad thematic area	**SUSTDEV**	**LIFESCIENCE**	**ICT**
Subprogramme in FP5	FP5-EESD	FP5-QOL	FP5-IST
Subprogramme in FP6	FP6-SUSTDEV	FP6-LSH	FP6-ICT
Scientific fields			
Biomedical Sciences	0.44	4.53	1.25
Basic Life Sciences	1.10	7.12	0.91
Biological Sciences	2.21	0.93	0.46
Chemistry and Chemical Engineering	2.20	0.73	1.36
Clinical Medicine	0.29	6.00	0.69
Computer Sciences	0.09	0.06	4.01
Earth Sciences and Technology	6.06	0.04	0.08
Environmental Sciences and Technology	6.17	0.32	0.09
Physics and Material Sciences	1.05	0.30	6.66
Electrical Engineering and Telecommunication	0.18	0.01	4.00

* only weights that are greater than one for at least one broad thematic area are shown

Table 2: List of variables included in Model A (acquisition of FP projects)

Dependent variables	**Description**
FP6 $LINKS_{ij}$	number of funded participations in 6^{th} FP between $region_i$ and $region_j$
Independent variables	
FP6 $WEIGHT_i$	number of funded participations of $region_i$ in 6^{th} FP
FP6 $WEIGHT_j$	number of funded participations of $region_j$ in 6^{th} FP
$COPUB_{ijt-3}$	number of international copublications between $region_i$ and $region_j$ in previous 3 years
FP5 RELATED LINKS	number of funded participations in related 5^{th} FP between $region_i$ and $region_j$
FP5 UNRELATED LINKS	number of funded participations in unrelated 5^{th} FP between $region_i$ and $region_j$
DISTANCE	straight line distance between $region_i$ and $region_j$
LANGUAGE	dummy the same language is spoken in $region_i$ and $region_j$

`

Table 3: Descriptive statistics and correlations for Model A (acquisition of FP projects)

	SUSTDEV	Mean	St. Dev.	Min	Max	1	2	3	4	5	6	7	8
1	FP6 $LINKS_{ij}$	1.022	2.818	0.000	65	1.000							
2	FP6 $WEIGHT_i$	4.511	1.923	0.000	7.847	0.325	1.000						
3	FP6 $WEIGHT_j$	4.442	1.974	0.000	7.847	0.340	-0.008	1.000					
4	$COPUB_{ijt-3}$	0.932	1.308	0.000	7.668	0.537	0.383	0.401	1.000				
5	FP5 RELATED LINKS	0.479	0.736	0.000	4.357	0.670	0.414	0.426	0.652	1.000			
6	FP5 UNRELATED LINKS	0.693	0.932	0.000	5.808	0.652	0.445	0.436	0.720	0.743	1.000		
7	DISTANCE	6.935	0.595	3.726	8.250	-0.059	-0.052	-0.076	-0.198	-0.079	-0.067	1.000	
8	LANGUAGE	0.963	0.190	0	1	-0.023	-0.021	-0.050	-0.154	-0.046	-0.050	0.295	1.000

	LIFESCIENCE	Mean	St. Dev.	Min	Max	1	2	3	4	5	6	7	8
1	FP6 $LINKS_{ij}$	0.623	2.315	0.000	96	1.000							
2	FP6 $WEIGHT_i$	3.434	2.326	0.000	7.591	0.293	1.000						
3	FP6 $WEIGHT_j$	3.569	2.304	0.000	7.591	0.280	-0.009	1.000					
4	$COPUB_{ijt-3}$	1.193	1.640	0.000	7.977	0.499	0.455	0.454	1.000				
5	FP5 RELATED LINKS	0.454	0.738	0.000	5.004	0.562	0.411	0.415	0.697	1.000			
6	FP5 UNRELATED LINKS	0.712	0.929	0.000	5.576	0.513	0.455	0.441	0.685	0.733	1.000		
7	DISTANCE	6.935	0.595	3.726	8.250	-0.089	-0.120	-0.086	-0.215	-0.060	-0.080	1.000	
8	LANGUAGE	0.963	0.190	0	1	-0.032	-0.048	-0.053	-0.181	-0.040	-0.054	0.295	1.000

	ICT	Mean	St. Dev.	Min	Max	1	2	3	4	5	6	7	8
1	FP6 $LINKS_{ij}$	1.073	3.363	0.000	129	1.000							
2	FP6 $LINKS_{ij}$	4.624	1.687	0.000	8.087	0.319	1.000						
3	FP6 $WEIGHT_i$	4.694	1.686	0.000	8.087	0.323	-0.008	1.000					
4	$COPUB_{ijt-3}$	1.032	1.461	0.000	7.683	0.497	0.415	0.425	1.000				
5	FP5 RELATED LINKS	0.436	0.739	0.000	5.384	0.689	0.447	0.449	0.653	1.000			
6	FP5 UNRELATED LINKS	0.725	0.928	0.000	5.384	0.567	0.458	0.464	0.684	0.729	1.000		
7	DISTANCE	6.935	0.595	3.726	8.250	-0.055	-0.079	-0.051	-0.200	-0.058	-0.079	1.000	
8	LANGUAGE	0.963	0.190	0	1	-0.025	-0.042	-0.050	-0.144	-0.046	-0.051	0.295	1.000

Table 4: Model A explaining acquisition of FP funding between region i and j

SUSTDEV		MODEL 1			MODEL 2	
	coef.	se	p-value	coef.	se	p-value
FP6 WEIGHT$_i$	0.995	0.007	0.000	0.953	0.009	0.000
FP6 WEIGHT$_j$	0.992	0.007	0.000	0.949	0.009	0.000
COPUB$_{ijt-3}$	0.024	0.005	0.000	0.010	0.006	0.093
FP5 RELATED LINKS				0.096	0.010	0.000
FP5 UNRELATED LINKS				-0.014	0.009	0.129
DISTANCE	-0.048	0.009	0.000	-0.052	0.009	0.000
LANGUAGE	-0.064	0.025	0.010	-0.075	0.025	0.003
CONSTANT	-10.741	0.081	0.000	-10.27	0.105	0.000
NUMBER OF OBS.	29679			29679		
LOG-LIKELIHOOD	-20869			-20825		

LIFESCIENCE		MODEL 1			MODEL 2	
	coef.	se	p-value	coef.	se	p-value
FP6 WEIGHT$_i$	1.018	0.008	0.000	1.007	0.009	0.000
FP6 WEIGHT$_j$	1.014	0.008	0.000	1.003	0.009	0.000
COPUB$_{ijt-3}$	0.004	0.005	0.411	-0.004	0.006	0.489
FP5 RELATED LINKS				0.040	0.011	0.000
FP5 UNRELATED LINKS				-0.011	0.009	0.225
DISTANCE	-0.031	0.010	0.003	-0.039	0.011	0.000
LANGUAGE	-0.028	0.027	0.300	-0.039	0.027	0.156
CONSTANT	-10.61	0.092	0.000	-10.44	0.110	0.000
NUMBER OF OBS.	29679			29679		
LOG-LIKELIHOOD	-13375			-13370		

ICT		MODEL 1			MODEL 2	
	coef.	se	p-value	coef.	se	p-value
FP6 WEIGHT$_i$	0.996	0.007	0.000	0.912	0.010	0.000
FP6 WEIGHT$_j$	1.004	0.007	0.000	0.921	0.010	0.000
COPUB$_{ijt-3}$	0.011	0.005	0.016	-0.005	0.005	0.325
FP5 RELATED LINKS				0.114	0.011	0.000
FP5 UNRELATED LINKS				0.011	0.009	0.233
DISTANCE	-0.062	0.010	0.000	-0.070	0.011	0.000
LANGUAGE	-0.044	0.024	0.070	-0.046	0.025	0.062
CONSTANT	-10.78	0.090	0.000	-9.863	0.129	0.000
NUMBER OF OBS.	29679			29679		
LOG-LIKELIHOOD	-21971			-21900		

Note: all regressions are estimated with robust standard erros

Table 5: List of variables included in Model B (effect of FP projects)

Dependent variables	**Description**
$COPUB_{ij}$	number of international copublications between $region_i$ and $region_j$
Independent variables	
main variable	
$FUND_{ijt-3}$	number of FP projects in which a regional pair has participated in previous 3 years (mean centered)
$COPUB_{ijt-3}$	number of copublications between $region_i$ and $region_j$ in previous 3 years (mean centered)
$FUND_{ijt-3}COPUB_{ijt-3}$	product of FUNDijt-3 (mean centered) and COPUBijt-3 (mean centered)
control variables	
PUB_i	number of publications of $region_i$ in previous 3 years
PUB_j	number of publications of $region_j$ in in previous 3 years
DISTANCE	geographic distance between $region_i$ and $region_j$
LANGUAGE	dummy whether the same language is spoken in a regional pair (0) or not (1)

Table 6: Descriptive statistics and correlations for Model B (effect of FP projects)

	SUSTDEV	Mean	St. Dev.	Min	Max	1	2	3	4	5	6	7	8
1	$COPUB_{ij}$	3.961	18.852	0.000	1134.053	1.000							
2	$FUND_{ijt-3}$	0.000	0.701	-0.438	3.931	0.369	1.000						
3	$COPUB_{ijt-3}$	0.000	1.358	-1.008	6.856	0.470	0.639	1.000					
4	$FUND_{ijt-3}COPUB_{ijt-3}$	0.608	1.592	-2.387	25.618	0.550	0.687	0.563	1.000				
5	PUB_i	8.563	1.552	0.860	12.555	0.199	0.396	0.471	0.233	1.000			
6	PUB_j	8.414	1.676	0.860	12.555	0.180	0.398	0.479	0.225	-0.005	1.000		
7	DISTANCE	6.935	0.595	3.726	8.250	-0.148	-0.079	-0.198	-0.086	-0.069	-0.011	1.000	
8	LANGUAGE	0.963	0.190	0.000	1.000	-0.198	-0.044	-0.164	-0.062	-0.024	0.028	0.295	1.000

	LIFESCIENCE	Mean	St. Dev.	Min	Max	1	2	3	4	5	6	7	8
1	$COPUB_{ij}$	10.872	51.664	0.000	1234.832	1.000							
2	$FUND_{ijt-3}$	0.000	0.630	-0.345	4.546	0.442	1.000						
3	$COPUB_{ijt-3}$	0.000	1.706	-1.302	6.746	0.480	0.669	1.000					
4	$FUND_{ijt-3}COPUB_{ijt-3}$	1.168	2.988	0.000	38.461	0.617	0.764	0.551	1.000				
5	PUB_i	9.358	1.737	3.224	13.722	0.219	0.405	0.498	0.265	1.000			
6	PUB_j	9.315	1.762	3.224	13.722	0.197	0.408	0.486	0.247	-0.013	1.000		
7	DISTANCE	6.935	0.595	3.726	8.250	-0.158	-0.083	-0.216	-0.091	-0.099	-0.045	1.000	
8	LANGUAGE	0.963	0.190	0.000	1.000	-0.216	-0.046	-0.191	-0.067	-0.043	0.000	0.295	1.000

	ICT	Mean	St. Dev.	Min	Max	1	2	3	4	5	6	7	8
1	$COPUB_{ij}$	6.038	28.903	0.000	1344.771	1.000							
2	$FUND_{ijt-3}$	0.000	0.655	-0.364	5.021	0.388	1.000						
3	$COPUB_{ijt-3}$	0.000	1.517	-1.123	6.965	0.471	0.621	1.000					
4	$FUND_{ijt-3}COPUB_{ijt-3}$	1.025	2.688	0.000	36.157	0.555	0.729	0.530	1.000				
5	PUB_i	8.847	1.704	2.331	13.193	0.205	0.388	0.485	0.231	1.000			
6	PUB_j	8.818	1.783	2.331	13.193	0.187	0.381	0.477	0.224	-0.006	1.000		
7	DISTANCE	6.935	0.595	3.726	8.250	-0.138	-0.059	-0.199	-0.077	-0.105	-0.017	1.000	
8	LANGUAGE	0.963	0.190	0.000	1.000	-0.167	-0.044	-0.153	-0.062	-0.038	0.032	0.295	1.000

Table 7:Model B explaining the number of co-publications after joint acquisition of FP-projects

SUSTDEV															
Model		**MODEL 1**			**MODEL 2**			**MODEL 3**			**MODEL 4**			**MODEL 5**	
Sample		$COPUB_{ij=0}$			$COPUB_{ij=0}$			ALL			ALL			ALL	
	coef.	**se**	**p-value**	**coef.**	**se**	**p-value**	**coef.**	**se**	**p-value**	**coef.**	**se**	**p-value**	**coef.**	**se**	**p-value**
$FUND_{ij}$				0.279	0.037	0.000	0.125	0.007	0.000	0.100	0.006	0.000	0.261	0.009	0.000
$COPUB_{ij}$										0.173	0.006	0.000	0.212	0.006	0.000
$FUND_{ij}COPUB_{ij}$													-0.074	0.003	0.000
PUB_i	0.819	0.015	0.000	0.790	0.016	0.000	0.841	0.006	0.000	0.716	0.007	0.000	0.698	0.007	0.000
PUB_j	0.793	0.014	0.000	0.767	0.015	0.000	0.832	0.005	0.000	0.710	0.007	0.000	0.688	0.007	0.000
DISTANCE	-0.466	0.030	0.000	-0.460	0.030	0.000	-0.409	0.009	0.000	-0.336	0.009	0.000	0.212	0.006	0.000
LANGUAGE	-1.368	0.106	0.000	-1.340	0.105	0.000	-0.617	0.023	0.000	-0.490	0.022	0.000	-0.336	0.009	0.000
CONSTANT	-10.455	0.267	0.000	-10.103	0.269	0.000	-12.450	0.100	0.000	-11.035	0.104	0.000	-10.665	0.105	0.000
NUMBER OF OBS.	71728			71728			148395			148395			148395		
LOG-LIKELIHOOD	-24142			-24113			-192086			-191716			-191413		

LIFESCIENCE															
Model		**MODEL 1**			**MODEL 2**			**MODEL 3**			**MODEL 4**			**MODEL 5**	
Sample		$COPUB_{ij=0}$			$COPUB_{ij=0}$			ALL			ALL			ALL	
	coef.	**se**	**p-value**	**coef.**	**se**	**p-value**	**coef.**	**se**	**p-value**	**coef.**	**se**	**p-value**	**coef.**	**se**	**p-value**
$FUND_{ij}$				0.342	0.037	0.000	0.231	0.007	0.000	0.150	0.007	0.000	0.351	0.011	0.000
$COPUB_{ij}$										0.226	0.004	0.000	0.247	0.004	0.000
$FUND_{ij}COPUB_{ij}$													-0.066	0.003	0.000
PUB_i	0.649	0.011	0.000	0.628	0.011	0.000	0.776	0.005	0.000	0.594	0.005	0.000	0.583	0.005	0.000
PUB_j	0.654	0.011	0.000	0.631	0.011	0.000	0.785	0.005	0.000	0.603	0.005	0.000	0.586	0.006	0.000
DISTANCE	-0.331	0.024	0.000	-0.332	0.023	0.000	-0.346	0.008	0.000	-0.242	0.008	0.000	-0.248	0.008	0.000
LANGUAGE	-1.071	0.076	0.000	-1.073	0.076	0.000	-0.772	0.021	0.000	-0.497	0.020	0.000	-0.491	0.020	0.000
CONSTANT	-11.574	0.222	0.000	-11.192	0.225	0.000	-13.721	0.096	0.000	-11.256	0.098	0.000	-10.935	0.100	0.000
NUMBER OF OBS.	71728			71728			148395			148395			148395		
LOG-LIKELIHOOD	-30219			-30179			-241356			-239929			-239631		

Table 7 (continued): Effect model explaining the # of co-publications after joint participation in FP-projects

ICT															
Model	**MODEL 1**			**MODEL 2**			**MODEL 3**			**MODEL 4**			**MODEL 5**		
Sample	$COPUB_{ij=0}$			$COPUB_{ij=0}$			ALL			ALL			ALL		
	coef.	**se**	**p-value**	**coef.**	**se**	**p-value**	**coef.**	**se**	**p-value**	**coef.**	**se**	**p-value**	**coef.**	**se**	**p-value**
$FUND_{ij}$				0.323	0.042	0.000	0.138	0.007	0.000	0.099	0.007	0.000	0.281	0.010	0.000
$COPUB_{ij}$										0.227	0.005	0.000	0.254	0.005	0.000
$FUND_{ij}COPUB_{ij}$													-0.069	0.003	0.000
PUB_i	0.679	0.013	0.000	0.656	0.013	0.000	0.747	0.005	0.000	0.598	0.006	0.000	0.580	0.006	0.000
PUB_j	0.680	0.013	0.000	0.658	0.013	0.000	0.747	0.005	0.000	0.598	0.006	0.000	0.579	0.006	0.000
DISTANCE	-0.395	0.030	0.000	-0.401	0.029	0.000	-0.308	0.010	0.000	-0.224	0.009	0.000	-0.232	0.009	0.000
LANGUAGE	-1.284	0.106	0.000	-1.252	0.105	0.000	-0.603	0.023	0.000	-0.434	0.021	0.000	-0.430	0.021	0.000
CONSTANT	-9.411	0.263	0.000	-9.060	0.266	0.000	-12.225	0.101	0.000	-10.445	0.099	0.000	-10.037	0.101	0.000
NUMBER OF OBS.	71198			71198			148395			148395			148395		
LOG-LIKELIHOOD	-25009			-24980						-212425			-212128		

Note: all regression include year dummies and are estimated with robust standard errors that are clustered on the regional pair level

Appendix A: We intend to post this appendix as a separate paper on the web, including the original data that is used for establishing the concordance tables.

A concordance table linking the Framework Programmes to scientific publications

Abstract: The objective of this short paper is to establish a concordance between various thematic areas within the fifth and sixth Framework Programme and scientific fields. Such a concordance is necessary to determine in which scientific fields we can reasonably expect acquisition and effect of Framework Programme funding to be concentrated. In order to establish the concordance we construct a unique database that covers approximately 10,000 scientific publications that are – at least partly – an outcome of FP funding. By linking these publications to various FP sub-programmes, we can characterise the FPs on the base of their scientific focus and scientific intensity.

Introduction

The EC's series of *Framework Programmes (FPs)* are specifically designed to pool resources and promote R&D collaboration in pre-competitive research between the EU member states. Their overarching goal is to improve the communication and collaboration among researchers and to boost Europe's innovation performance. Several decades of pursued policies are likely to have had a substantial influence on the European scientific landscape. Though the EU national governments still provide the bulk of the European research expenditure, the influence of the FPs might well be extensive due to: (a) matching requirements by national governments or organisations which alone may double the volume of funding; (b) the agenda setting role of European programs; (c) the quality of EU funded research which may result in multiplier effects on domestically funded research, especially in the less advanced countries, and (d) continuation of collaborations after funding halts as networks tend to be self-sustaining.

However, not much is known about the actual output that results from participation in the Framework Programmes. This makes it difficult to keep track of the potential effect of FP funding on the European scientific and technological landscape. The objective of this paper is therefore to report on a database that contains scientific output of the FPs which allows us to establish a concordance between various subprogrammes within the FPs and the cognitive focus of their output (i.e. journals, scientific fields). In doing so, we rely on a unique database that covers approximately 10,000 scientific publications that are - at least partly – an outcome of funding within the Framework Programmes. Using unique grant numbers and call abbreviations of the FPs available in the *the EUPRO database* (Roediger-Schluga and Barber

2006), we develop a search algorithm that links publications to the various subprogrammes in the Fifth and Sixth Framework Programme (referred to as FP5 and FP6 from now on). We subsequently proceed to characterize these programmes in terms of scientific intensity, and scientific focus

The focus of the analysis is on thematic areas in the FP5 and FP6. The rationale for this focus is threefold: (i) the thematic areas receive the lion share of funding within the FPs. More specifically, in FP5 they mutually make up 72,5% of funding, whereas in FP6 63,3% of funding is allocated within those priorities, (ii) one of the major goals of the thematic areas is their scientific and technological impact, implying that scientific publications can be considered a significant and meaningful output of the funded projects[9] (iii) knowledge production within the thematic areas explicitly concentrates on collaborative actions which allows for a sound comparison not only with regional nodes but also with regional links.

Methodology

We draw from all citation databases[10] available on the internet as part of Thomson Reuters' "Web of Science" (WoS) (http://portal.isiknowledge.com), in order to obtain a set of publications that are funded within the FPs. As of February 2009, the indexed publications in the various citation databases can be searched on grant activity (i.e. funding agencies and grant numbers) and funding acknowledgements, going back to mid-2008. This allows for identifying a set of publications that results from funded research within particular

[9] Publications in peer-reviewed journals are possibly an indicator on which Framework Programme projects will be monitored in the future.
[10] Available citation database are the expanded version of the Science Citation Index, the Social Sciences Citation Index and the Arts & Humanities Citation Index.

Framework Programme projects. The identification, retrieval and structuring of such a set of publications is explained below in four steps.

Step 1: Identifying FP-funded publications in Web of Science

Publications that are at least partly funded by the FPs are expected to list the contributing European project, program, institutions as a funding agency in their publication. We therefore developed a query that searches Web of Science using the field tag "Funding Agency (FO)"[11] for names and abbreviations related to European institutions, EC RTD policies and the FPs. An understanding of the most frequently used funding-agency-terms was - next to common sense - obtained by using the 'analyze results' tool within Web of Science. This web-based tool lists the most frequently occurring funding agencies on a set of publications resulting from a particular query. Using queries with random samples of publications (e.g. FO=1, FO=a, etc.) we checked whether all terms related to European institutions and European RTD policies were included in our query. This resulted in the following query:

FO= (European Com* OR European Un* OR EU OR EC OR FP* OR Framework P* OR European project OR European Network OR European Research Area) (1)

Although the various citation databases list funding agencies back to halfway 2008, we limited the query to the year 2009 (i.e. "PY=2009") to cover an equal and full year of journal

[11] Within Web of Science one can also use the field tag "Funding Text (FT)". Yet, this text lists the full acknowledgement text that can be found on the original publication, including all kinds of personal and institutional acknowledgements that do not result from direct funding of the project. Hence, we solely focus on funding agencies to prevent a large number of false positives (publications that are found but are not funded within FPs).

issues for all indexed journals. The search was performed on February 19, 2010 and yielded a total of 23,415 records.

Step 2: Structuring funding agencies and grant numbers

All 23,415 records were downloaded and structured in such a way that every row indicates a publication record with additional information listed in five columns: (i) Unique number (UT), (ii) Journal (SO), (iii) Journal Category (SC), (iv) Funding Agency/Grant number (FO/FG), and (v) Funding Text (FT). On the base of the information in the column Funding Agency/Grant number (FO/FG), publications can be attributed to a corresponding Framework Programme and project, whereas the information in the columns Journal and Journal Category make it possible to aggregate groups of funded publications to scientific fields.

Often, a publication is funded by multiple agencies and lists more than one funding agency and grant number. In the WoS indexed publications, this listing of funding agencies and grant numbers follows a particular format in which funding agencies and their assigned grant numbers are linked (funding agency [grant number]; funding agency [grant number]). Based on this format, we use query (1) to separate EU funding agencies and grant numbers (from now on EU funding text) from non EU funding agencies and grant numbers (from now on non EU funding text).

The rationale for doing so is twofold. First, the EU funding text is a 'clean' representation of funding provided by European institutions. Hence, a query within those records to identify specific FP projects and thematic areas is very unlikely to pick up false positives (i.e.

publications not funded by the European Union but identified as such). This makes it redundant to check every record manually for correct assignment.

However, the funding acknowledgement texts are often long text strings that are directly taken from journal articles and for which the conversion to standardized fields within Web of Science is not always straightforward. This is for instance because funding agencies and their grant numbers are not always directly combined in the acknowledgement text, or because the number of funding agencies and the number of grant numbers do not exactly correspond. It occurs as a consequence that grant numbers of European projects are linked to non-EU funding agencies and vice versa. Indeed, we do not want to exclude those records, but we also cannot automatically be sure that they result from funding within the FPs (i.e. there may be other institutions that assign similar grant numbers). Thus, we also need to check the non-EU funding texts to identify records that otherwise would have become false negatives (i.e. publications that are funded by the European Union but that are not identified as such).

Step 3: Identifying funding agencies and grant numbers within publications

The EUPRO database provides information on unique grant numbers that are assigned to each FP project. In order to link funded publications to the projects and thematic areas by which they are funded we build an algorithm that searches for parts of grant numbers within the two funding texts. A code is assigned to a publication record, every time a predefined search term is found. As the identified publication records stem from the year 2009, it is very unlikely that we will identify projects that are funded before 1998. Hence, we chose to only focus on projects funded within FP5 (1998-2002) and FP6 (2002-2006).

The developed coding schemes to search for funded publications are based on the structure of the grant numbers and on names of the thematic areas. With respect to the structure of the grant numbers, a 'model case' grant number in FP5 starts with a four digit code that contains both a three digit text and a one digit number. Together the first four digits correspond to a specific program, sub-program or call of an FP5 thematic area. Grant numbers in FP5 are subsequently followed by the term 'CT', the starting year of funding and a non-unique five digit number (e.g. ENK5-CT-2000-00305, QLK3-CT-2002-30237). In FP6, grant numbers have a similar structure although the first part of the grant number is not necessarily standardised in three text digits and one number digit and the last part of the grant number is a unique six digit number instead of a non unique five digit number (e.g. FOOD-CT-2005-534684). Moreover, the first part of the FP6 grant numbers is sometimes omitted in the publication records because of the unique character of the latter six digit code.

With this information in mind, we program a coding algorithm that searches for the names and grant numbers of all distinct thematic areas in FP5 and FP6. A publication record is assigned a 1 if either the code at the start of the grant number (e.g. ENK5, NMP1) or a name related to the specific program (e.g. Information Society Technologies), is mentioned in the EU funding text. A code of 2 is assigned to a publication record if any of the above terms is found in the non-EU funding text. In case of FP6 a publication record is also assigned a 1 or a 2 if the unique six digit grant number is found in the EU or non EU funding texts. Furthermore, in case of FP5 we assign a code of 3 (for EU funding text) or a code of 4 (for non EU funding text) to publication records that do not list the entire unique four digit term but only the unique three digit text at the start of the grant number.

As mentioned, the number of false positives in publication records with a code of 1 is expected to be negligible because the search range is very narrow (i.e. only EU funding texts) and because the search terms are non-generic and unique across thematic areas. The number of false positives in publication records that are coded 2 is expected to be higher as some of the search terms may also be used by non EU funding agencies. Hence, publication records that are coded 2 are always manually checked on their correct assignment to a thematic area[12]. Naturally, FP5 publications with a code of 3 or 4 are also always checked manually. Manual checks are always performed on the entire funding text to make sure that erroneously separated funding agencies and grant numbers could be detected. In case such a check did not resolve the issue we traced back the entire funding acknowledgement text to get a better understanding of the funding context of the research. In case we could not judge from the acknowledgement text we did not assign the publication to the thematic area. The final search procedure for all thematic areas is summarized in Table 1.

After running this procedure, 8,235 publication records could be assigned to a specific thematic area within FP5 or FP6. It turns out that in most thematic areas scientific publications are a frequently occurring output and even the thematic areas of FP5 still render almost 1,400 publications in 2009. Three programmes stand out for their low publication intensity: FP5-GROWTH, FP6-AEROSPACE, FP6-CITIZENS. In case of the former two this may well be due to a relative focus on technological development rather than scientific advancement which is confirmed by an extensive participation of private actors in these

[12] The thematic area Information Society Technologies and corresponding terms such as IST are both present in FP5 and FP6, while the thematic area New Materials Production and corresponding terms such as NMP are both present in FP6 and FP7. In the former case, we used a slightly different coding scheme and publication records for FP5 and FP6 were merged, whereas in the latter case we did an additional manual check based on the year of funding that is listed in the grant numbers.

programmes[13]. In case of the latter, low publication intensity may be explained by the social-scientific character of FP6-CITIZENS, because we know that codification processes including the listing of funding acknowledgements are less common in the social sciences than in the natural sciences. For these reasons we decide to exclude those three thematic areas from further analysis.

The total matching rate is approximately 35% of the initially 23,415 identified publications. The matching rate is not very high because we defined the initial query in a rather general way. This implies that a significant part of the initially identified 23,415 publication records is funded by the European Union, but not specifically within the thematic areas of FP5 and FP6. Examples include publications funded within FP4 or FP7, publications funded within non-thematic areas of FP5 and FP6 (e.g. Marie Curie Training Network, Research Infrastructures), publications funded by other European institutes or programs (e.g. European Regional Development Fund, European Structural Fund) and publications that list an institute of the European Community as a funding source but do not specify any program or grant number.

The concordance

All publication records contain information on the journal in which they are published. WoS also aggregates journals in journal categories based on their titles and citation relations (see for an overview and critique: Leydesdorff and Rafols 2009). Indeed, every publication belongs to one journal, but the journals are often attributed to more than one journal category

[13] An analysis of organisations that participate in FP5-GROWTH and FP6-AEROSPACE reveals indeed that participation shares of private actors are of all thematic areas the highest for FP5-GROWTH (47,9%) and FP6-AEROSPACE (45,9%).

(with a maximum of 6). On their turn, the journal categories can be uniquely assigned to one out of 35 scientific fields as defined by the CWTS/NOWT classification (Tijssen et al. 2010).

This structure allows us to aggregate all funded publications of a thematic area to journal categories and scientific fields. In so doing, a fractional counting method is used in which every publication is counted as one. Accordingly, the journal categories or scientific fields to which publications are assigned receive a fractional count equal to one divided by the total number of journal categories or scientific fields to which the publication is assigned[14]. We do not include the social science journals and journal categories in the analysis as the coverage of social sciences in Web of Science is relatively low and the listing of funding acknowledgements is not very common.

For every thematic area we end up with a distribution of publications across journal categories and scientific fields. The cognitive focus of the various FPs is particularly visible at the level of scientific fields. Figure 1 and Figure 2 show the distribution of FP output across scientific fields. In all thematic areas more than 50 percent of publications are produced in two or three scientific fields.

In order to assess the overlap between different thematic areas in FP5 and FP6, we also correlate thematic areas on their distribution of publications across journals, journal categories and scientific fields. The results are shown in Table 2. It becomes clear that the programmes FP5-EESD and FP6-SUSTDEV and the programmes FP5-QOL and FP6-LSH are particularly correlated, especially at the level of scientific fields.

[14] A scientific field is always counted once even if a publication is assigned to more than one journal category belonging to the same field.

Conclusion

In this short paper we developed a concordance between various subprogrammes of the FPs and their distribution of output across scientific fields. Based on this concordance we can reasonably infer in which scientific fields participation and effect of the FPs are concentrated. We showed that publications are a frequently occurring output for FP projects. There are also high correlations between the scientific output of related programmes in FP5 and FP6, suggesting patterns of continuity in cognitive focus of the programme. In a follow-up analysis we intend to relate publication output to commonly used bibliometric indicators, including co-authorship networks and citation counts.

References

Leydesdorff, L., Rafols, I. (2009) A global map of science based on the ISI subject categories. *Journal of the American Society for Information Science and Technology*, 60: 348-362.

Roediger-Schluga, T., Barber, M. (2006) The structure of R&D collaborations networks in the European Framework Programmes: data prcessing, network construction and selected results. *International Journal of Foresight Innovation Policy*, 4: 321-347.

Tijssen, R.J.W., Hollanders, H., Kanerva, M., van Leeuwen, T.N., Nederhof, A.J., Van den Berg, P. (2010) Wetenschaps en Technologie-indicatoren 2010, Netherlands Observatorium van Wetenschap en Technologie (NOWT), *Report for the Netherlands Ministry of Education, Science and Culture*.

Table 1: Search procedure

	5th Framework Programme			6th Framework Programme						Both
	EESD	GROWTH	QOL	AERO	CITIZ	FOOD	LSH	NMP	SUSTDEV	IST
MATCHED PUBLICATIONS	431	86	865	86	6	726	2816	1106	1028	1085
NAMES										
Programme name (incl. variants)	▭ ✋	▭ ✋	▭ ✋	▭ ✋	▭ ✋	▭ ✋	▭ ✋	▭ ✋	✋ ✋	▭ ✋
GRANT NUMBER FP5										
Combination of first 3 digits and year										▭ ✋
First 4 digits	▭ ✋	▭ ✋	▭ ✋							
First 3 digits	✋ ✋	✋ ✋	✋ ✋							✋ ✋
GRANT NUMBER FP6										
First 4 digits				▭ ✋	▭ ✋	▭ ✋	▭ ✋	▭ ✋	▭ ✋	▭ ✋
Last 6 digits				▭ ✋	▭ ✋	▭ ✋	▭ ✋	▭ ✋	▭ ✋	▭ ✋
Less than 6 digit with brackets/hypens										✋ ✋
OTHER										
Years of programme								✋ ✋		

▭ **and** ✋ **indicate automatic and manual search in EU funding text (first sign) and non EU funding text (second sign)**

Table 2a: Correlations between different thematic areas of the Framework Programmes

JOURNALS

	FP5-EESD		FP5-QOL		FP6-FOOD		FP6-LSH		FP6-SUSTDE		FP6-NMP		FP-IST	
	r	p-value	*r*	p-value	*r*	p-value	*r*	p-value	*r*	p-value	*r*	p-value	*r*	p-value
FP5-EESD	1.00													
FP5-QOL	-0.01	0.67	1.00											
FP6-FOOD	-0.02	0.29	0.26	0.00	1.00									
FP6-LSH	-0.07	0.00	0.35	0.00	0.06	0.00	1.00							
FP6-SUSTDEV	0.47	0.00	0.01	0.73	0.00	0.83	-0.07	0.00	1.00					
FP6-NMP	-0.03	0.21	-0.06	0.01	-0.04	0.07	0.01	0.74	0.09	0.00	1.00			
FP-IST	-0.05	0.01	-0.06	0.00	-0.05	0.01	0.03	0.18	0.02	0.44	0.50	0.00	1.00	
N	2193		2193		2193		2193		2193		2193		2193	

JOURNAL CATEGORIES

	FP5-EESD		FP5-QOL		FP6-FOOD		FP6-LSH		FP6-SUSTDE		FP6-NMP		FP-IST	
	r	p-value	*r*	p-value	*r*	p-value	*r*	p-value	*r*	p-value	*r*	p-value	*r*	p-value
FP5-EESD	1.00													
FP5-QOL	0.17	0.02	1.00											
FP6-FOOD	0.08	0.25	0.63	0.00	1.00									
FP6-LSH	-0.06	0.43	0.76	0.00	0.35	0.00	1.00							
FP6-SUSTDEV	0.91	0.00	0.14	0.05	0.06	0.39	-0.04	0.57	1.00					
FP6-NMP	-0.01	0.88	-0.04	0.53	-0.02	0.77	0.04	0.54	0.13	0.07	1.00			
FP-IST	-0.06	0.42	-0.05	0.47	-0.06	0.38	0.05	0.45	-0.01	0.92	0.46	0.00	1.00	

Table 2b: Correlations between different thematic areas of the Framework Programmes

SCIENTIFIC FIELDS

	FP5-EESD		FP5-QOL		FP6-FOOD		FP6-LSH		FP6-SUSTDEV		FP6-NMP		FP-IST	
	r	p-value	*r*	p-value	*r*	p-value	*r*	p-value	*r*	p-value	*r*	p-value	*r*	p-value
FP5-EESD	1.00													
FP5-QOL	0.02	0.93	1.00											
FP6-FOOD	-0.01	0.96	0.82	0.00	1.00									
FP6-LSH	-0.10	0.64	0.88	0.00	0.71	0.00	1.00							
FP6-SUSTDEV	0.95	0.00	0.02	0.92	-0.02	0.94	-0.05	0.82	1.00					
FP6-NMP	-0.05	0.83	-0.08	0.73	0.01	0.96	0.04	0.87	0.16	0.47	1.00			
FP-IST	-0.17	0.46	-0.10	0.65	-0.12	0.60	0.00	0.99	-0.08	0.72	0.64	0.00	1.00	
N	22		22		22		22		22		22		22	

Figure 1: Distribution of publications across scientific fields for FP5

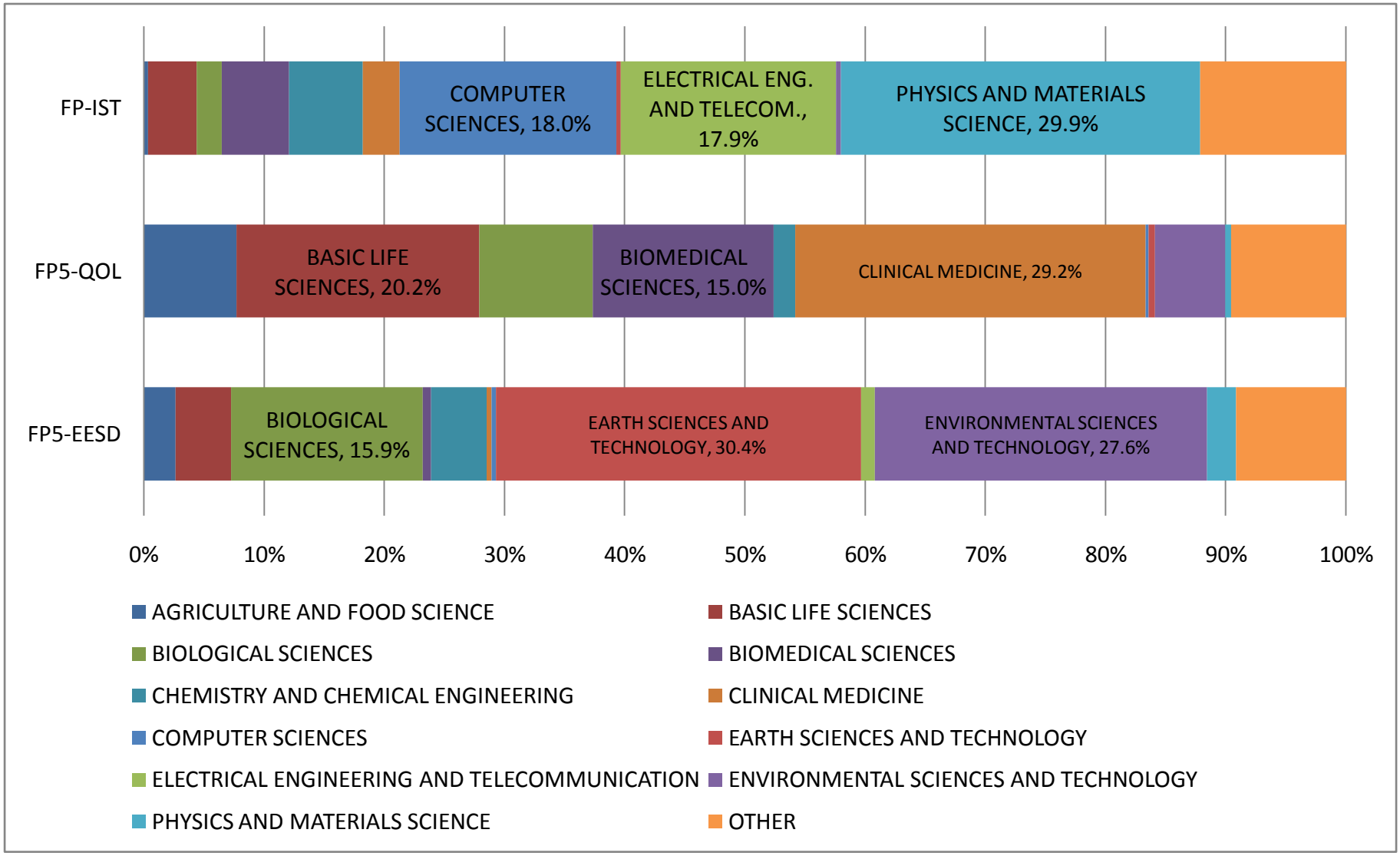


Figure 2: Distribution of publications across scientific fields for FP6