Evaluating a Research Funding Program: Measuring the Impact of EU 6th Framework Programme

Ed Noyons[1], Robin Teigland[2], Andrew Schenkel[2], and Jan Maier[3]

[1] *noyons@cwts.leidenuniv.nl*
Centre for Science and Technology Studies (CWTS), Leiden University, Leiden (The Netherlands)

[2] *robin.teigland@hhs.se; andrew.schenkel@hhs.se*
Stockholm School of Economics (Sweden)

[3] *j.maier@avedas.com*
Avedas AG (Germany)

Abstract

The European Union, as well as national, regional, and even local government organizations across Europe, is promoting research and development collaborations as efficient and effective vehicles to turn ideas into new products, services, and solutions in order to improve competitiveness. One of the largest programs is the European Community Sixth Framework Programme (FP6) for Research, Technological Development and Demonstration (RTD). This paper investigates the impact of the FP6 funding using two output measures developed for this study: organizational network structure and publication impact. Results using a dataset consisting of 1312 RTD projects with more than 14,000 partner organizations indicate that 1) different types of organizations play different roles in the networks, 2) RTD partner co-publication was impacted positively, and 3) the size of the RTD project (i.e., number of partners) influences project performance.

Introduction

The European Union, as well as national, regional, and even local government organizations across Europe, is promoting research and development collaborations as efficient and effective vehicles to turn ideas into new products, services, and solutions in order to improve competitiveness. Within recent years, the European Union has funded numerous R&D programs with the purpose of bringing together organizations in related fields from across the EU member states to pursue research and technology development. One of the largest and most recent programs is the European Community Framework Programme 6 (FP6) for Research, Technological Development and Demonstration (RTD). The two overarching goals of FP6 according to the European Commission are the following: 1) strengthening the scientific and technological bases of industry and encourage its international competitiveness while 2) promoting research activities in support of other EU policies.[5] The FP6 has a budget of 17.5 billion Euros, and funding from the program started in 2003 and ended in 2006, with some projects continuing until 2010. During the FP6, the EU funded approximately 2500 projects or RTD collaborations in the fields of Life Sciences, Information Science Technologies, Nanotechnologies, Food Quality and Safety, and Citizens and Governance alone.

To fulfil the overarching goals of the EU 6th Framework Program, two primary objectives were developed:1) the development of a European Research Area (ERA) through improved integration and co-ordination of research and 2) the strengthening of the competitiveness of the European economy. More specifically, FP6 stated: 'It aims at scientific excellence, improved competitiveness and innovation through the promotion of increased co-operation, greater complementarity and improved co-ordination between relevant actors, at all levels'. The main underlying assumption of FP6 is that research in Europe is largely fragmented, and

[5] http://ec.europa.eu/research/fp6/pdf/fp6-in-brief_en.pdf

Proceedings of ISSI 2009, pages 313–324.
Edited by B. Larsen and J. Leta.

competitiveness requires improved integration and coordination. To accomplish this, projects have been structured into 7 thematic RTD areas. In addition, several different program instruments, or ways of stimulating collaboration through various project constellations, were implemented. The program instruments are listed below, together with their most important objectives and features.

Integrated Projetcts (IP)	• Multi-partner projects • To support objective-driven research • Primary deliverable is knowledge
Networks of Excellence (NoE)	• Multi-partner projects • Aimed at strengthening scientific and technological excellence
Specific Targeted Research Projects (STREP)	• Multi-partner research, demonstration or innovation projects. • Purpose is to support research etc, (more limited scope and ambition than IPs)
Coordinated Action (CA)	• Multi-partner actions • Intended to promote and support the networking and coordination of research etc.
Specific Support Action (SSA)	'... contribute actively to the implementation of activities of the work programme, the analysis and dissemination of results or the preparation of future activities ...'

Finally, each RTD thematic area consists of projects, composed by a number of member organizations. These organizations can consist of SMEs and large firms, research institutes, academic hospitals, universities, as well as other public sector organizations.

The return from the investment of billions of Euros into science research as a result of such a funding program have been studied frequently, and from a science policy point of view, the results of such studies are of great importance (c.f., Luukkonen, 1998, Laredo, 1998 and Molas-Gallart & Davies, 2006). All policymakers want to bet on the right horse. In all such studies the key question is what were the expected outcomes of the funding program? In other words, what was the money to be used for? Only with this understanding are we able to make a proper evaluation of the impact of the funding program. In line with this, we set out to investigate the impact of FP6 on collaboration among research actors in Europe. More specifically, we examined 1312 RTD projects including more than 14,000 participating organizations within five thematic research areas and four program instruments to determine the impact of FP6 funding and thus to evaluate the effectiveness of FP6. Below we present first our method and data, our results, and then our conclusions. For as far as we know this is the first ex-post evaluation of a large funding program in which on such a scale quantitative and qualitative has been integrated.

Data and method

In total, FP6 included RTD areas consisting of 10,065 projects with 73,552 participants and 42,196 unique participants. The comprehensiveness of these data meant that we had to limit the number of RTD areas that form the basis for exploring our research question. Thus, from the nineteen RTD areas we selected five areas in agreement with the European Commission based upon the following criteria.

1. Life sciences, genomics and biotechnology for health (LSH)
2. Information society technologies (IST)

3. Nanotechnologies and nanosciences, knowledge-based multifunctional materials and new production processes and devices (NANO)
4. Food quality and safety (FOOD)
5. Social sciences: Citizens and governance in a knowledge-based society (CITI)

All FP6 project data were then provided by the European Commission (DG Research). Firstly, we required high quality data. This meant that we needed to identify for each RTD: 1) the participating RTD member organizations and 2) the names and addresses of the principal investigator (PI) from each RTD partner organization. Secondly, we wanted to increase diversity in terms of the number, type, and location of RTD member organizations within each RTD such that comparisons between thematic areas could be made. In total, these five thematic areas represented 25% of all FP6 RTD projects, 43% of all RTD participants and 38% of all unique RTD participants. Table 1 shows descriptive data of the selected RTD thematic areas.

Table 1: Total number of projects started in 2003, 2004, and 2005

FP6 Thematic Area	Number of projects started in 2003-2005	Total number of FP6 projects	Percentage
Life sciences genomics and biotechnology for health (LSH)	311	599	52%
Information society technologies (IST)	524	1093	48%
Nanotechnologies and nanoscience NMP)	325	446	73%
Food quality and safety (FOOD)	95	190	50%
Citizens and social sciences (CITIZENS)	57	146	39%
Total in five FP6 areas	1312	2474	53%

Data cleaning and RTD project selection

The project dataset was cleaned thoroughly, with the experience from preceding projects and information from the Web of Science as far as relevant. A combination of sophisticated software tools and manual cleaning yielded a considerable and vital improvement of the data quality. As an illustration of the effect of our cleaning activities, we included Figure 1 below. It shows the decrease of unique names of partners in FP6 projects by thematic area.

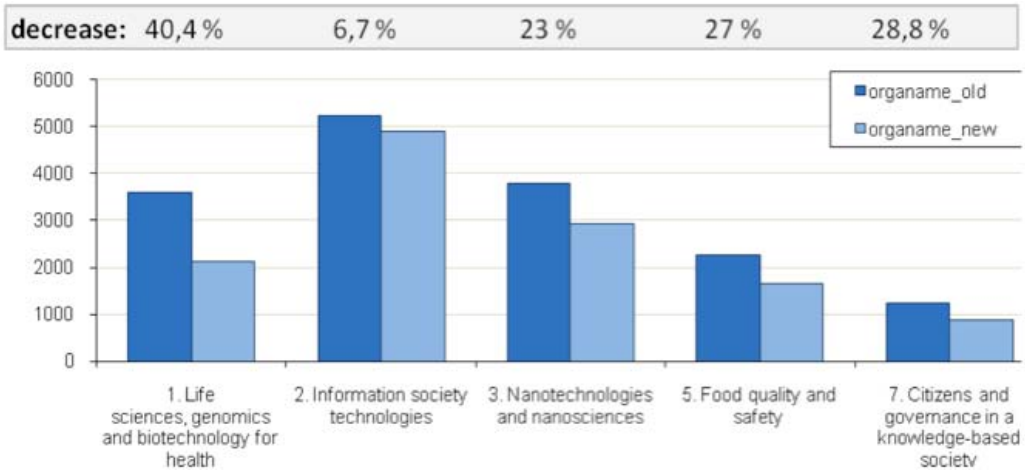

Figure 1: Unique values for organization names

Due to this thorough cleaning activity, we were able to reallocate a considerable part of partners initially classified as 'Other' to more specific types such as 'Industry'. In the network analysis this has led to a much more accurate picture of the role of industry in the field of IST. In this area the number of industry partners increased with no less than 50%.

To strengthen our ability to draw conclusions, we decided to delimit the set of projects so that we could obtain a zero measurement from which data and trends associated with each RTD project could be made. As the projects started at different times during the period of FP6, we used the following time classification for the respective RTD projects in order to treat the projects in a consistent manner with respect to time:

- Period 0: the four years preceding the RTD project starting month plus 1 year after the project starting month
- Period 1: the period starting one year after the starting month of the RTD project to the end of 2007

This method of treating projects meant that projects that started after 2005 were excluded from the dataset since we would not fully capture period 0 for all projects, making comparisons or trends difficult to observe. From our initial selection of the five thematic areas, the number of RTD projects selected of the total number of projects started in 2003, 2004 and 2005 are in the below table. These 1312 projects included more than 14,000 partners.

Table 2: Total number of projects started in 2003, 2004 and 2005

RTD Thematic Area	Number of projects started in 2003-2005	Total number of FP6 projects	Percentage
Life sciences genomics and biotechnology for health	311	599	52%
Information society technologies	524	1093	48%
Nanotechnologies and nanoscience	325	446	73%
Food quality and safety	95	190	50%
Social sciences	57	146	39%
Total in Five Areas	1312	2474	53%

Defining outcome measures

Our first step was to identify a suitable output measure that could be used across thematic areas and instruments. This proved to be a difficult step as there are no data in the FP6 database that may be used to directly represent performance. However, for the first time in the framework program, principle investigators (PIs) are registered in the project data for FP6. As such we are now able to create a link, albeit a tenuous one, between RTD projects and publications. Our idea is to then create a set of performance measures for each RTD project that reflects the publications associated with each of the principle investigators (PI) in the RTD. Before we go into detail about our method and findings, we would like to mention that there are some limitations to this method since clearly we cannot state with confidence that the publications by a PI truly are a direct result of his/her participation in the RTD**.** To be more specific, a publication may have been in the publication pipeline prior to the start of the RTD; however, due to lag in various stages of this pipeline, the publication may appear after the RTD had started or even ended. Or the publication may even be the result of the PI's participation in another project and not in the FP6 RTD. However, our assumption is that some of the publications from PIs are the results of their RTD work and this would then be picked up in the publication data.

However, to attempt to adjust for this inability to create a direct link between RTD

participation and publication, we go one step further and investigate the number of co-publications by PIs in an RTD. In other words, using the Web of Science bibliometric database, we were able to identify publications that were co-authored by two or more PIs from the same RTD. In this manner, the likelihood that the publication is a result of the RTD is higher. While some of these publications may be the result of PIs collaborating prior to the RTD, they are however more likely to be directly related to the purpose of the RTD with the FP6 continuing then to fund their collaborative endeavors. Finally, one limitation with this method is that we are not able to capture all the publications that are authored by individuals in the RTD who are not PIs for their organization since we do not have the names of all the participating members from each organization in the RTD. Having stated some of the limitations of the method, below we provide a more detailed view of our method.

In line with the above, we collected only publications by the principle investigators (PI) from the Web of Science (WoS), a multi-disciplinary bibliography of more than 6000 international peer reviewed journals. We consider the output of the entire organisation as irrelevant because they represent work done remote from the research within the FP6 project context. For each PI involved in the 1312 projects, we collected his/her publications in the Web of Science in the period 1998-present. Furthermore, we tagged each publication collected for partners in a project as belonging to T0 or T1.

- T0 is the period from 4 years before the start of a project until one year after the start,
- T1 is the period from one year after the start of a project until present.

Moreover, we identified those papers in which more than one partner in the same project collaborated as co-authors. These Partner Co-publications (PcP) in T1 are considered to be the best proxy of the effect of funding. It turns out that we were able to collect a significant amount of corresponding publications in the first four thematic areas. However, we were not able to collect sufficient data for an analysis of the Social Sciences area. Thus, the focus of our study presented here will be on the first four areas.

To harvest the publication data for each PI, we developed a sophisticated search strategy. Although this approach is uniquely developed for this project, it has similar principles as studies discussed in Smalheiser and Torvik (2009) and as Costas and Bordons (2007). We started with a search in the Web of Science using surname and first initial of the PI combined with an address (partner organisation). Obviously, this search strategy allows many false positives, due to the ambiguous author names and due to the lack of strict search strings (surname plus right truncated first initial). Moreover, in specific cases we added alternative names if required. This is applied in the case of authors with multi part surnames, e.g., Spanish names. In all cases, we calculated the similarities between all retrieved publications, based on their patterns with respect to:

- shared co-authors
- shared addresses
- shared cited references.

By applying a weight to each of these elements we were very well able to define disjoint clusters in each set of papers assigned to a PI. We compiled a decision rule to assign a set of publications (a cluster) to the PI or not. We used additional PI information, such as e-mail address and full name to establish this. Thus we were able to rule out the false positives and thus to retrieve reliable sets of publications belonging to a PI. Moreover, we were able to deal with the problem of ambiguity in general and of synonymous author names. In the network below (Figure 2), we depicted the example of a Spanish PI, Maria Bermejo-Sanz.

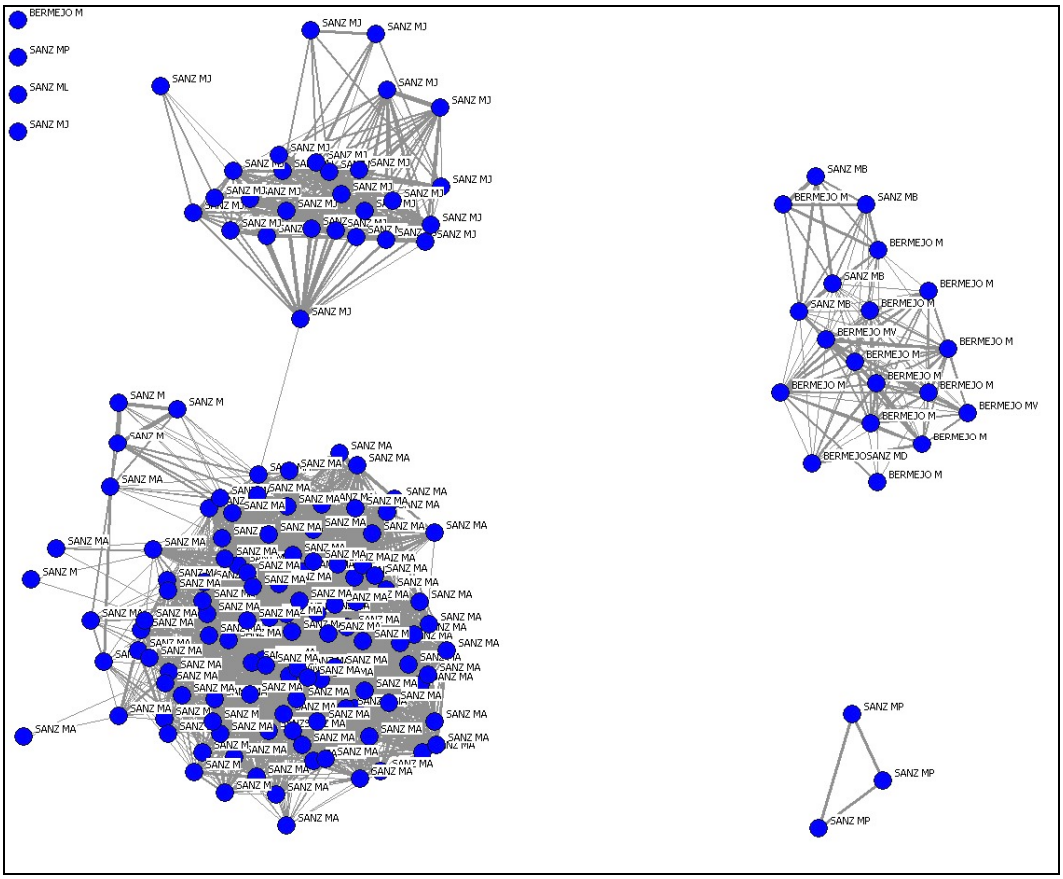

Figure 2: Publication clustering of PI *Maria Bermejo-Sanz*

In this example, we depict (pajek network analysis) the retrieved publications of a PI *Maria Bermejo-Sanz*. We searched for her publications using *Bermejo* M*'* and *Sanz M**. Each publication in the network was labelled by the name variant on the basis of which it was retrieved. The labelling of the nodes in the different clusters shows the strength of the method, as the cluster on the upper-right hand side contains all relevant publications of this PI, while the different labels shows the variants on the basis of which they were retrieved. The other clusters evidently contain the false positives and are ruled out.

Main results

Below we discuss the networks funded by the FP6 program regarding general morphologies and particular positions of roles of actors. Subsequently we will discuss the impact PF6 seems to have had on the publication activity of participants. And finally, we will link up by discussing the results we found in relation to the outcomes of the survey among participants. This survey was held among some hundreds of FP6 participants randomly chosen by area and instrument. They were involved through email and filled out the questionaire through a dedicated website. In total for 95% of the 133 selected projects we received input from participants, covering 32% of the selected organisations (747).

Networks of project data

The first part of the analysis concerns a thorough network analysis of the project data to characterize and visualize the collaboration in projects at different levels of aggregation. Thus, we reveal the structure as funded by the FP6 program. This includes the network morphology as such but also the roles of the most prominent actors (the organisations) and of the different types (e.g., higher education vs. industry).

As the different instruments have different objectives, we would expect the networks of the individual instruments to have their own specific characteristics, i.e., their own specific morphology. We describe the morphology using two features: clustering and centralization. Clustering means the existence of a critical mass of connected actors who are specialized in certain research topics. If there are a lot of organizations that cooperate in more than one project we regard a network as more clustered and stable, the other way around its structure would be more hierarchical. Clustered networks are more stable, hierarchies are more efficient. Centralization refers to the question if a network is single peaked or multi peaked. A single peaked network is dominated by one or a few very central actors also linked with each other (this is also called hub assortativity). Multi-peaked networks consist of more than one centre. They have a higher innovation potential because more peaks mean that there are groups of competing actors with different resources (e.g. knowledge) and competing new ideas. We determine centralization by comparing the distributions of the values that represent the actor's interaction intensity.

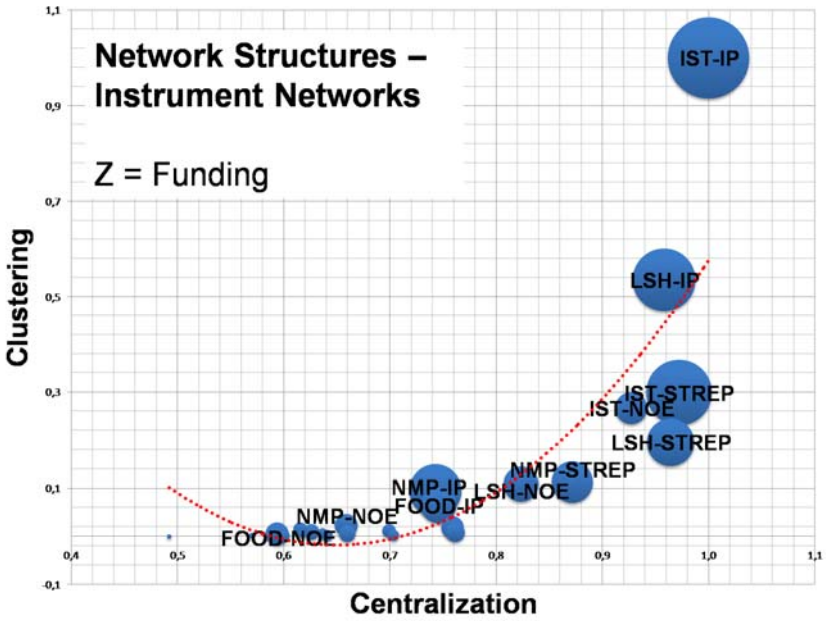

Figure 3: Network Morphology of Instrument Networks

We summarize the morphology of the instruments as follows. In Figure 3, we depict the instruments on the basis of their centralization (horizontal axis) and clustering (vertical axis), while the size of a circle indicates the amount of funding spent within the different instruments in LSH.

Without going into detail of the results of this analysis, we note that the different instruments in the different areas show large differences. However, the differences cannot be explained from the different focus of the instruments. More research into the morphological aspects is needed to elaborate on this aspect.

With respect to the roles of the different sectors, we created a network per thematic area. In Figure 4 below, we depicted as an illustration the position and roles of the five main types of organisations (universities, research institutes, hospitals, and industry) and a rest category (other).

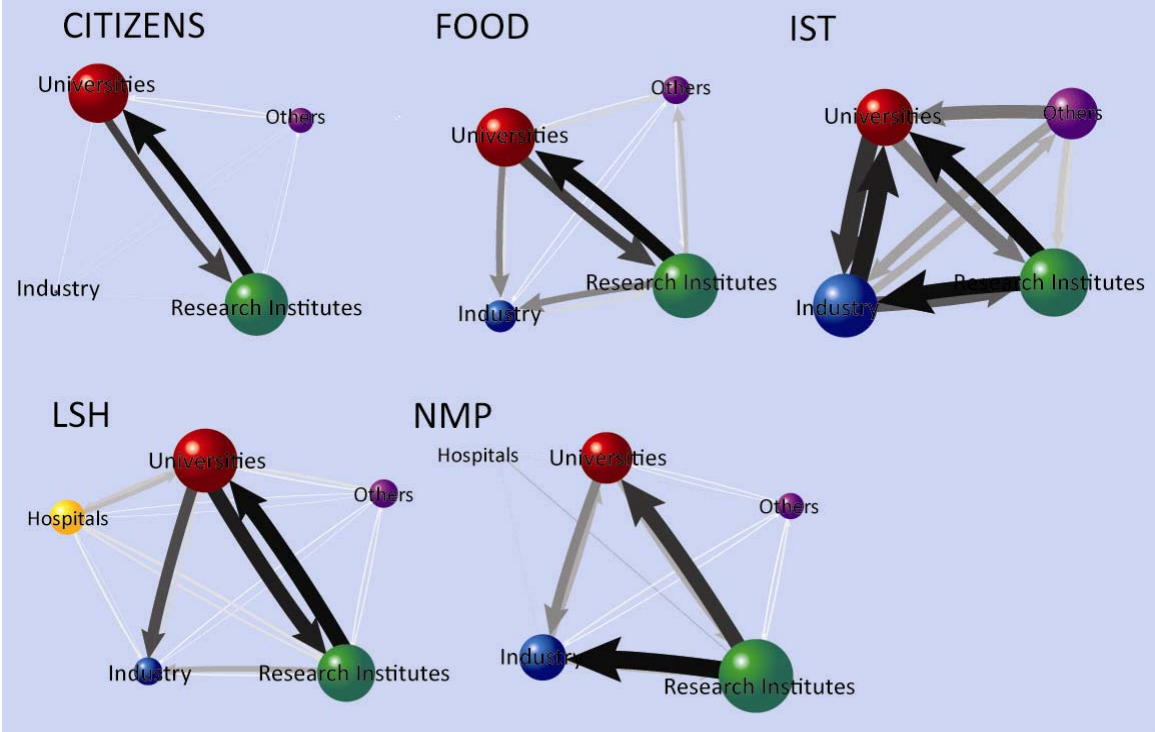

Figure 4: Coordination network in 5 thematic areas as funded by FP6

In these visualizations, circles represent institution types, and arcs indicate coordination. An arc leading from the universities to industry means that there is at least one project which is coordinated by a university and in which an industry actor participates – the stronger and darker an arc the higher the number of projects. We can clearly see that in LSH universities and research institutions "coordinate each other" (symmetrical linkages), and companies mainly just participate (in projects which are coordinated by universities). In IST coordination is an activity of industry, universities, and research similarly. And all three types are connected through reciprocal relationships. These results suggest that roles, strategies, and expectations of companies (and of the other types) are different in LSH and IST.

Impact on publication activity

The networks on the basis of project data do not represent the impact of the funding as such. It is the outcome of the review process of FP6. In that sense we can consider this as the input network. How this funding strategy has impact on the research area and to meet the objective of improving collaboration will be investigated by measuring partner co-publication (PcP) activity. First, some general statistics: from the over 14,000 partners (i.e., Principle Investigators, PI) we were able to collect publication data for more than 6000. The distribution of publications by field is given Table 3 below.

Table 3: Numbers of PI publications by field

Field	T0	T1
Food	7,621	4,033
IST	8,493	5,685
LSH	28,388	14,213
Nano	17,016	8,568

Note that T0 and T1 do not represent a fixed period in time but are different time periods for each individual project. T0 is the time from 4 years before the start until one year after the

start. T1 is the period from one year after the start of a project until ultimately 6 years after start (but in all cases until present). In that sense the output as such does not say anything on the impact of the program. Only that in both periods there is a significant amount of papers to base our analysis upon. A second step in our bibliometric approach concerns the identification of co-publications of partners in a project. These partner co-publications (PcPs) are considered to be the best proxy of the scientific output related to an FP6 project. In Figure 5 below, we depicted the share of PcP in T0 and T1 for the 4 investigated thematic areas.

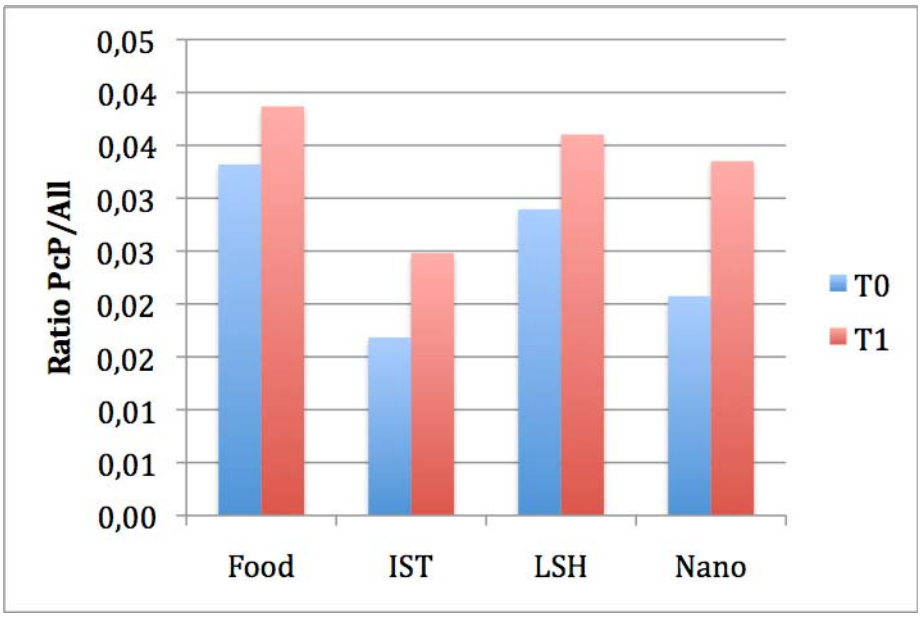

Figure 5: Share of PcP in T0 and T1 by thematic area

Although the numbers of PcP's form a small fraction of the total output of PI's, it is interesting to see that the proportion increases significantly in all areas. We may conclude that the thematic areas represent such different research fields that this is a general trend, that cannot be explained from the fact that, in general, researchers tend to increase their output as covered by the Web of Science overall in the past ten years. Note that we have depicted proportions and not absolute numbers.

Our next approach concerns the different instruments. By distributing the publication output over the instruments by thematic area, we get a good view on their impact on science output. Figure 6 below shows the results.

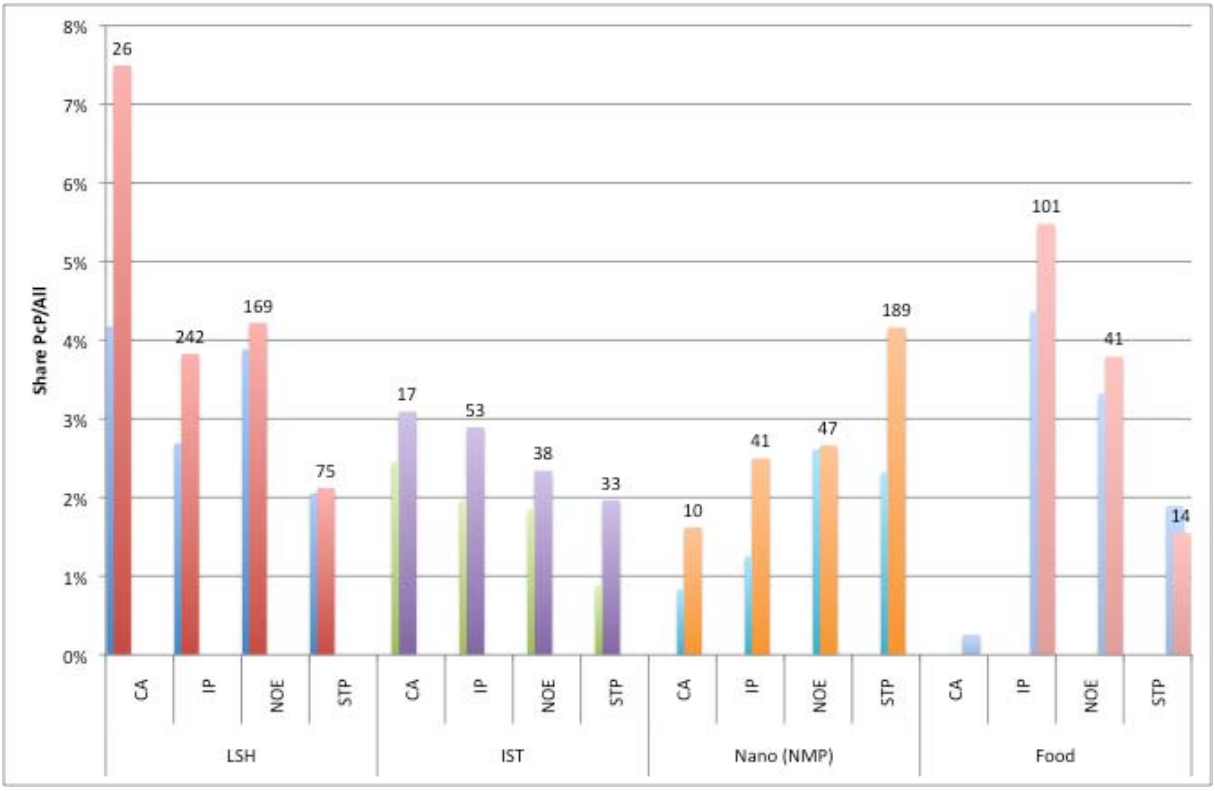

Figure 6: PcP proportion by instrument, by thematic area

From this figure we may conclude that in all fields, IP (integrated projects) has a similar impact in terms of increased PcP. Also in NoE (networks of Excellence) we measure this impact in all fields. The impact of STREP is particularly measured in IST and Nano (NMP). CA had no significant impact. The numbers are also too low in all four fields. It should be noted that both in IP as well as in NoE, the PcP activity was already measured in T0, so that we may conclude that the impact of FP6 was to enhance the existing network. This is particularly the case for NoE.

Finally, we investigated whether there was any measurable effect from the size of a project team. Hereto, we identified three classes of teams: large (more than 10 participants), medium (less than 10 but more than 5) and small (less than 5). For these classes we calculated the average proportion of PcP in T0 and T1.

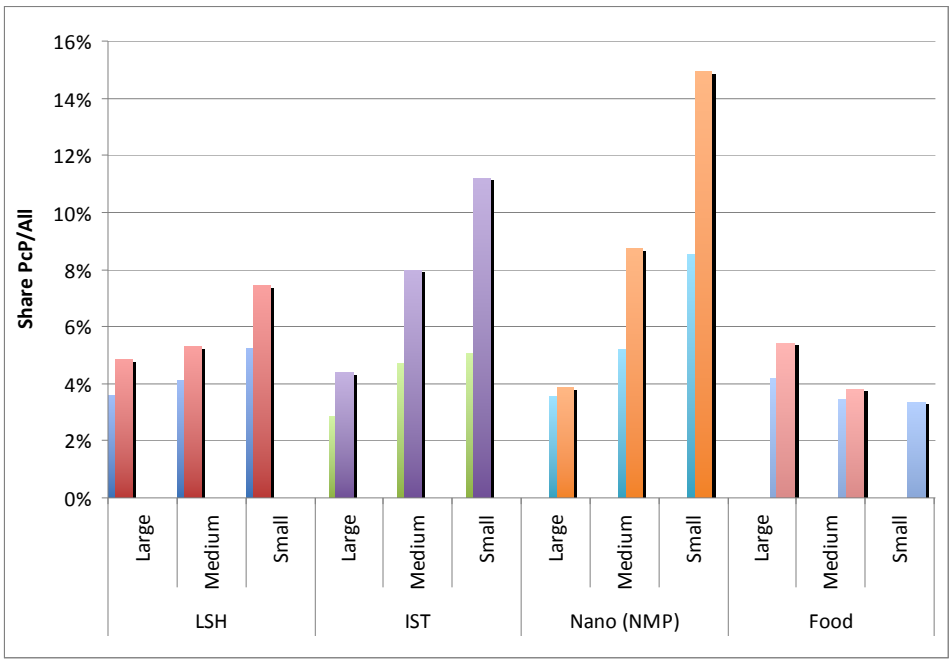

Figure 7: Share of PcP by field, by team size

The results in Figure 7 show an increase in all classes in T1 as compared to T0, which coroborates the earlier results. In addition, these results show that particularly the small teams (5 partners or less) not only have on average more PcPs but they also show a higher increase in T1 as compared to T0. Only in the area of Food were the amounts too small (only 52 projects with a very high percentage of large teams) to contribute to the overall findings. In the next section, we will discuss a finding from the survey which validates this overall finding.

Survey

Because we realize that behind the statistics there are stories of people working in a project, we conducted a survey of 747 organizations in 133 RTD projects (32% response rate). Partly, the survey was meant to check the validity of our analysis and indicators and partly we were looking for impacts that were not measurable with the data available. Since the number of organizations and RTDs was in the thousands and we were limited in our ability to conduct a widescale survey, we decided to select RTDs based on the dissimilar case selection method (Yin, 1989). In this manner, we identified the RTDs that were at opposite sides of the identified scales of interest in order to facilitate the ability to generalize our findings. To select the RTDs and their member organizations for the survey based on this method, we identified RTDs based on two measures of performance: publication impact and RTD project centrality (since this is generally accepted as an indicator of performance), within each instrument and thematic area. To be more specific, for example, for the thematic area of IST, we identified for each instrument a) the two RTDs that had the highest co-publication impact, b) the two RTDs that had the lowest co-publication impact, c) the two RTDs that had the highest project centrality in the FP6 project network, and d) the two RTDs that had the lowest project centrality in the FP6 project network. We then constructed an online survey to which an invitation to complete the survey was sent in an email to the principle investigator for each of the organizations identified. In this email, there was a hyperlink to the online survey, and participants had two weeks to complete the survey. Two reminder emails were sent. The survey required approximately 30 minutes to complete and questions were grouped around the topics of rationale for FP6 project collaboration, project dynamics, and project outcomes.

The first notable result was that the respondents indicated that the most relevant outcomes of their FP6 project was publications, methods and collaboration, rather than commercial products, patents and the like. This validates our approach to measure the impact of FP6 by using publication data. If anything measurable, it will be by counting publications and co-publications.

Another important finding is the effectiveness of the different instruments and the network characteristics as a whole on the impact of the program. The results of the survey showed that the more participants in a project, the less effective the project was with regard to financial planning and management as well as to meeting the objectives of the project. Furthermore, the results of the survey showed that the collaboration in a project was more efficient if partners knew each other before the start of the project or even had collaborated before. These outcomes confirm the outcomes of FP6, particularly for IP and NoE, being a platform to enhance existing networks of collaboration.

Conclusion

Our analysis shows that the FP6 funding created a structure with different sectors, i.e., government, industry, academia, playing different roles in different thematic areas. We consider this an important issue because the EC has acknowledged that the role of, for instance, industry should not be underestimated in certain areas. In addition, FP6 funding has led to an improved network as measured by the partner collaboration in terms of publication output. With respect to the instruments, we found that particularly Integrated Projects (IP) have had a positive impact on collaboration. It appears that through this instrument the partner co-publication proportion has increased. However, in the case of NoE, we found a more modest increase, perhaps because the funding went to already existing networks of researchers. And although the aspect of being already acquainted is an important issue to partners in a project to be more efficient, we should be careful not to allow these networks to become too big because the survey also revealed that the bigger the networks, the less efficient they become. This means that there seems to be some kind of optimum team size with respect to effectiveness.

Acknowledgments

This study has been funded by the European Commission through the project: 'Structuring Effects of Community Research – The Impact of the Framework Programme on Research and Technological Development (RTD) on Network Formation' (30-CE-0130222/00-22).

References

Costas, R and Bordons, M (2007). Algorithms to solve the lack of normalization in author names in bibliometric studies. *Investigacion Bibliotecologica*, 21, 13-32.

Laredo, P (1998). The networks promoted by the framework programme and the questions they raise about its formulation and implementation. *Research Policy*, 27. 589–598.

Luukkonen, T (1998). The difficulties in assessing the impact of EU framework Programmes. *Research Policy*, 27. 599–610.

Molas-Gallart, J and Davies, A (2006). Toward Theory-Led Evaluation: The Experience of European Science, Technology, and Innovation Policies. *American Journal of Evaluation*, 27. 64-82.

Smalheiser, NR and Torvik VI (2009). Author name disambiguation. *Annual Review of Information Science and Technology*, 43. 287-313.

Yin, R.K. 1989. Case Study Research: Design and Methods. Newbury Park: Sage.