Who gets Horizon 2020 research grants? Propensity to apply and probability to succeed in a two-step analysis

Simen G. Enger[*,1,2] and Fulvio Castellacci[2]

1) Centre for Technology, Innovation and Culture, University of Oslo, P.O. Box 1108 Blindern, 0317 Oslo, Norway
2) Norwegian Ministry of Education and Research, Department of Research, P.O. Box 8119, 0032 Oslo, Norway

* Corresponding author: simen.enger@tik.uio.no; sien@kd.dep.no

Accepted version

Please cite the article as follows:

Enger, S.G. & Castellacci, F. (2016). Who gets Horizon 2020 research grants? Propensity to apply and probability to succeed in a two-step analysis. *Scientometrics.* doi:10.1007/s11192-016-2145-5

Abstract

This paper presents a timely analysis of participation in the 8th European Framework Programme for Research and Innovation (EU FP) Horizon 2020. Our dataset comprises the entire population of research organizations in Norway, enabling us to distinguish between non-applicants, non-successful applicants, and successful participants. We find it important to distinguish two stages of the participation process: the self-selection stage in which organizations decide whether they wish to apply for EU funding, and the second stage in which the European Commission selects the best applications for funding. Our econometric results indicate that the propensity to apply is enhanced by prior participation in EU FPs and the existence of complementary national funding schemes; further, that the probability of succeeding is strengthened by prior participation as well as the scientific reputation of the applicant organization.

Keywords: Horizon 2020; EU Framework Programs; research funding; research policy; higher education institutions; public research organizations.

Acknowledgements
We would like to thank two anonymous referees of this journal for valuable comments and suggestions. The authors are solely responsible for any remaining error and omission. Work on this paper was supported by the Norwegian Research Council's *Public Sector Ph.D. Program* [grant number 246964/H20].

1. Introduction

Since the first European Framework Programme (EU FP) was established in 1984 with the objective of strengthening scientific and technological collaboration in Europe, its importance has increased steadily (Breschi et al. 2009; Ortega and Aguillo 2010). The budget has grown from just below 4 billion Euros in the first framework program (FP1) to almost 80 billion Euros for the 8th and current Horizon 2020. With the establishment of the European Research Council in 2007, EU FPs have become a central source of funding for applied and basic research (Nedeva 2013).

With the increasing importance of EU research funding, national policymakers in European countries have put domestic participation in EU research at the center of the research policy agenda. In Norway, for instance, the government has explicitly stressed greater participation as an essential part of national strategies for internationalization of research. As of early 2015, Horizon 2020 had received 36,000 applications from all EU/EEA member-states; funding was granted to slightly less than 5,000 of these. Norwegian researchers were involved in 1,530 applications, but only 216 were granted funding. That amounts to 1.79% of the available competitive funds – still below the national goal of at least 2% (Norwegian Ministry of Education and Research 2014). For national policymakers, Norwegians and others alike, the need to understand what determines participation in EU FP, and how this can be strengthened, ranks high on the agenda.

Academic research has recently started to investigate the factors that affect participation in EU FPs. One of the factors noted in this research relates to the formation of collaborative networks and the enduring nature of these (e.g. Defazio et al. 2009; Protogerou et al. 2010). Some research organizations have been found to participate repeatedly in EU-funded research, and function as central nodes in EU FP networks (Paier and Scherngell 2011; Roediger-Schluga and Barber 2008; Protogerou et al. 2010).

Among the factors that have been investigated to explain these patterns, scientific capabilities of applying institutions (in particular, their academic reputation and scientific productivity) appear crucial (Geuna 1998; Lepori et al. 2015; Nokkala et al. 2011; Henriques et al. 2009). Another important dimension concerns the characteristics and structure of national funding, which may have complementarity or substitutability effects with international funding from the EU (Luukkonen and Nedeva 2010)

Due to limitation in terms of data availability, the literature thus far has analyzed the determinants of EU FP participation by focusing solely on the sub-sample of actual applicants, ignoring all the research organizations that decided not to apply for funding to EU FPs in the first place (e.g. because of lack of interest, time and/or resources). However, the latter group is relevant for this type of analysis. Knowing more about the research organizations that decide not to apply to EU FP may provide new insights on the underlying motivations for applying and the related obstructing factors.

Therefore, in order to advance our understanding of what determines participation in EU-funded research, we find it important to consider two distinct stages of this participation process. The first is the self-selection process, where some organizations decide to apply for funding, while many others decide not to. The second stage is the selection process carried out by authorities of the European Commission (EC), at the end of which some of the applicants are successful and are granted funding for one or more EU projects, whereas most other applicants are not. By distinguishing these two stages of the participation process, we seek to analyze the extent to which the factors highlighted in previous research have differing effects on the two stages of the participation process.

Our empirical analysis covers the entire population of Norwegian research organizations, using data on all 1,402 applications submitted by Norwegian research institutions to Horizon 2020 (hereafter: H2020) between 2014 and until early 2015. Aggregating these data at the organizational level, we match them with detailed national R&D statistics for the full population of public research organizations (PROs) and higher education institutions (HEI) in Norway, as well as registry data on whether they participated in FP6 (2003–2006), FP7 (2007–2013), and/or received national funding from the Research Council of Norway, RCN (2013–early 2015).

The empirical results show that the factors highlighted in the literature matter, but they do so differently in the two stages of the participation process: the propensity to apply to H2020 is enhanced by prior participation in EU FPs and by complementary national funding schemes; the probability of success in obtaining funding is strengthened by prior participation as well as the scientific reputation of the applying organization.

This paper is organized as follows: section 2 reviews the relevant literature on EU FP participation, and it points out our theoretical framework and hypotheses; section 3 presents the empirical context, data and indicators; section 4 discusses the econometric results; and section 5 summarizes the main findings and policy implications.

2. Theory and hypotheses

2.1 The literature on participation in European framework programs

Participation in EU FPs has come to rank high on national political agendas, motivating researchers to seek to understand the determinants and impact of EU research programs. The academic literature ranges from policy-oriented papers and evaluations to academic studies of the effect of EU FP participation on scientific productivity and innovation, as well as the organizational-level determinants of participation.

Much of the literature has focused on the establishment of networks and the collaborative structures underlying participation in EU FPs (Hoekman et al. 2012; Paier and Scherngell 2011; Breschi and Cusmano 2004; Protogerou et al. 2010; Defazio et al. 2009; Breschi et al. 2009; Roediger-Schluga and Barber 2008). This is because EU FP programs typically require collaboration between different research organizations and users in order for projects to be considered eligible for funding.

A part of the literature on collaborative networks has analyzed the additionality effects, in terms of increased scientific and innovative output, resulting from participation other than the formation of networks as such (Polt and Streicher 2005; Luukkonen 2000, 1998). Focusing on private companies, Matt et al. (2012) argue that EU FP participants are unlikely to contribute to radically new scientific knowledge, but are more likely to support networks in exploratory research than if the organizations form networks outside the EU frame. Defazio et al. (2009) hold that EU FP funding itself has a more direct effect on research productivity rather collaboration within the network. They found that it was only after the funding period had ended that collaboration affected the level of productivity.

Other studies, like Di Cagno et al. (2014), have held that EU FP participation has an effect on the transfer of knowledge and R&D spillovers. One case study of university participation echoes the results from company-firm oriented studies: according to Primeri and Reale (2012), the main output from participation was scientific papers, not commercial products or processes. The main argument for participation was the opportunity to collaborate with colleagues abroad, but also contribute to new scientific discoveries.

A common result in studies of collaborative networks in EU FPs is the observation of continued persistence of some organizations that hold central roles in different networks continuously over time, indicating that prior participation matters for successful project applications (Protogerou et al. 2010; Paier and Scherngell 2011; Roediger-Schluga and Barber 2008; Breschi and Cusmano 2004; Godø et al. 2009; Okubo and Zitt 2004; Makkonen and Mitze 2016). In particular, large, highly reputed European institutions seem to dominate as regards participation (Annerberg et al. 2010). Makkonen and Mitze (2016) argue that there exist well established oligarchic networks within the EU. Hence, even with new EU member-state participating in cross-country research, there are still strong networks among the large and older member-states. This has let several to suggest cumulative effects or a "Matthew effect" for organizations with already dominant positions in the research landscape (Protogerou et al. 2010).

The factors behind participation in EU research are complex and should in principle be investigated in a multi-level setting, simultaneously taking into account factors at the country level (characteristics of national science policy and research funding), the organizational level (strategies and funding characteristics of universities, faculties, research institutes, and departments), and the individual level (e.g. the capability and experience of the individual researchers applying for EU funding). All three levels of analysis are important and interact in complex ways. In practice, however, research has often focused on one level of analysis and neglected the others, depending on data availability and the specific objective of each study. Two sets of factors in particular have been discussed as possible determinants of EU FP participation, and for explaining the persistence and cumulative effects noted above.

One part of the literature has focused on country-level factors: the composition of national research systems and the national funding structure as a dimension affecting successful participations in EU FPs, investigating in particular whether there is commensurability between national and EU R&D policies (Dinges and Lepori 2006; Luukkonen and Nedeva 2010). National R&D schemes with a high degree of international orientation are said to have positive effects for building researcher capacity in applications to the EU (Dinges and Lepori 2006). Whether national policies are converging towards EU policies and if this promotes participation in EU FPs has been studied, as in the case of France, (Laredo 1998), Norway (Langfeldt et al. 2012) and Finland (Hakala et al. 2002). The conclusions are basically the same: internationalization of national R&D schemes leads to greater mobilization and participation in EU-funded projects. For example, Langfeldt et al. (2012) found that for parts of the 6^{th} and the 7^{th} EU FP, where calls were closely related to Norwegian R&D policies and schemes, a high degree of participation was evident. By contrast, Lepori et al. (2015) and Geuna (1998) controlled for country effects, and found only find slight evidence of it affecting EU FP participation for European HEIs. Others (Okubo and Zitt 2004; Tijssen 2008) have argued that small countries (e.g. Ireland) are more oriented towards research collaboration in EU than larger countries (e.g. Germany) which tend to cooperate domestically or outside the EU. However, other researchers find that the patterns of small state–EU collaboration are far less homogeneous (Ukrainski et al. 2014).

Another part of the literature has focused on organization-level factors, emphasizing the importance of organizational characteristics and capabilities as key determinants of participation and success in EU FPs (Hakala et al. 2002). Drawing on the literature of cumulative mechanisms (Viner et al. 2004), Geuna (1996) have suggested that the participation process is driven by reputation of the organization. The few studies to deal with factors at the organizational level (Lepori et al. 2015; Geuna 1998, 1996; Nokkala et al. 2011) have shown

that the distribution of participation is indeed skewed: a few universities account for most of the participation in EU funding in each country. These also tend to be among the top universities in Europe (Henriques et al. 2009). A key determinant, according to these studies, concerns the institution's level of scientific productivity – in terms of number of publications per full-time equivalents (FTE) (measure of productivity) and number of citations per publication (measure of reputation, or scientific impact). Also other organizational-level factors appear to matter, especially the size of the applying organization, and its scientific field(s) (Lepori et al. 2015; Hakala et al. 2002; Geuna 1998). At the national level, Ukrainski et al. (2014, p. 854), argue that countries with high shares in overall scientific output collaborate more widely within Europe.

2.2 Hypotheses

Previous studies of the determinants of EU FP participation have focused on the organizational-level and country-specific factors that can explain why some project applications receive funding whereas others do not. These studies have typically made use of databases of organizations applying to EU FPs, often linked to other national data sources on the characteristics and capabilities of research organizations (e.g. Geuna 1998; Lepori et al. 2015). We follow the same general approach, focusing on research organizations as the main unit of analysis in the theoretical framework and empirical model. The decision to participate in EU-funded research, and the capacity to do so, entail complex processes that should ideally be investigated in a multi-level setting, studying the interactions between country-, organization- and individual-level characteristics. However, data on individual-level characteristics (such as research experience, the reputation of individual participants in EU applications) are often not available. For that reason, we have chosen to focus on public research organizations as the main unit of analysis.

One main issue with the standard empirical approach previously used in this field is that, due to the lack of relevant data, it has concentrated on those who actually applied for EU FPs funding, ignoring all other research organizations that decided not to apply. We hold, however, that, in order to understand what determines participation in EU-funded research, it is important to consider two distinct stages of this participation process. The first is the self-selection process, whereby some organizations decide to apply for EU funding while many others decide not to do so. The second stage is the selection process, in which some applicants are successful and are granted funding for one or more EU projects, whereas most other applicants are not. In other words, at any time, the population of research organizations in a given country can be divided into three distinct groups as regards EU funding: (1) non-applicants, (2) unsuccessful applicants, (3) successful applicants. The literature to date has focused on the second and third group, whereas we have had access to data for the first group, and this enables us to distinguish the two above-mentioned stages in the selection process.

We do not intend to introduce new explanatory factors, but seek to investigate the extent to which the factors highlighted in previous research have differing effects on the two stages of the participation process. Our empirical analysis takes into account many of the variables and control factors already investigated in the literature, but it focuses on the three explanatory dimensions emphasized in recent works: (1) whether the organization has previously participated in an EU FP project (indicating persistence and cumulativeness effects); (2) national funding characteristics (e.g. complementarities between national and EU funding schemes and R&D policy); and (3) organizational research capabilities (e.g. reputation and scientific productivity). We put forward

three pairs of hypotheses, each focusing on one of these explanatory dimensions; further, we indicate whether they are expected to be more relevant in the first stage of the participation process, in the second, or in both.

First, we examine the role of prior participation in EU FPs. In line with the literature, we see this as a fundamental factor in strengthening the capacity of research organizations to get EU-funded projects. We argue that previous participation in FPs is an important determinant for both stages of the participation process, although for different reasons. In the first (self-selection) stage, when research organizations consider whether or not to invest time and effort in developing an EU project application, they make an assessment of the amount of time and resources they would have to invest in working on an application, and they compare these costs with the (uncertain) benefits that could accrue from participation. A first-time application to EU FP will entail a series of *sunk costs:* fixed costs (investments) that the organization must sustain in order to be able to apply for EU funding that first time – but these are costs the organization will not have to incur for subsequent future applications. Two types of sunk costs are relevant here. The first is related to the need to gather information and knowledge on the application procedure: the organization must build up specific administrative capacity for dealing with EU research. The second type of costs is instead related to network and team building. When a research organization wants to apply for EU funding for the first time, a major challenge involves finding suitable and competitive European partners in its field of research. This is no easy task, as the persistent and often closed nature of EU research networks makes it hard for new partners to be admitted into existing competitive networks. By contrast, an organization that has already participated in the recent past will not incur in these sunk costs: it already has in place the necessary administrative capacity and international network, making the prospect of working on a new application less costly and more attractive.

Once the decision to apply has been made, cost-benefit considerations will no longer be relevant, and the selection process will be affected primarily by the quality of the international project team and its application. These factors, in turn, are arguably strengthened by prior participation, since an organization that has taken part in an EU-funded collaborative project in the recent past has surely benefitted from *learning effects* concerning the application procedure, EU research priorities, and the quality of potential partners. These learning effects are likely to strengthen the organization's ability to submit a high-quality proposal and hence obtain funding. We summarize these arguments in the following hypotheses.

H1.1: *Previous participation in FPs increases the probability that an organization will apply for H2020 funding*.

H1.2: *Previous participation in FPs increases the probability that an organization will succeed in getting H2020-funded projects*.

Second, we shift the focus to national funding characteristics. As noted, the capacity of a research organization to participate in EU FPs may be affected by the characteristics and structure of national funding, especially whether there exist complementarities between national and EU funding schemes, and the degree of international orientation of national science and research policy. While agreeing with previous analyses on this point, we hold that the structure of national funding matters in the *first* stage of the participation process, not in the second.

Clearly, the availability of research funding serves as a primary motivation in the self-selection stage where organizations decide whether or not to apply to EU FP. On the one hand, with plentiful external funding to be obtained from domestic sources (e.g. RCN projects) there is less need to seek funds elsewhere. On the other hand, however, the availability of external funding from national sources may have complementarity effects, increasing the pool of resources that can be invested for international project applications. If the national authorities focus on the importance of EU programs in seeking to promote the internationalization of the national research system (as is the case in Norway and several other countries today), political priorities, discourse and specific financial support schemes may make a research organization more willing to invest time and resources in developing an EU project application.

However, these considerations are arguably not relevant for the second stage of the participation process. Once an organization has made a cost-benefit analysis and decided to submit a project application, the probability of success will not be directly affected by the structure and characteristics of national funding. The peer-review selection procedure conducted by the EC focuses on the quality and relevance of the project team and the proposal submitted, and it seems reasonable to expect the characteristics of national R&D policy to have little influence here. In principle, it could be argued that the experience with project acquisition and management gained by participating in national funding schemes can bring learning effects and perhaps foster the ability to obtain EU funds as well. However, that argument does not seem relevant for the specific case considered in this paper, as the procedures for project application and management in Norway (especially as regards the Research Council of Norway) differ from those adopted by the EC for H2020 projects. We therefore disregard the possible existence of learning and complementarity effects in the second stage of the participation process, and formulate our hypotheses as follows:

H2.1: *Funding granted by national institutions increases the probability that an organization will apply for H2020 funding*.

H2.2: *Funding granted by national institutions does not increase the probability that an organization will succeed in getting H2020-funded projects*.

Third, there is the role of organizational research capabilities, a factor that of increasing importance. It concerns two specific aspects: the scientific reputation of an organization (e.g. as measured by its citations and impact factor) and its scientific productivity (e.g. the number of recent publications or publication points). While finding this dimension generally relevant, we hold that research capabilities will matter differently in the two stages of the participation process: they will be largely irrelevant in the self-selection stage (or even go in the opposite direction from that indicated in earlier studies), but will be highly relevant in the second stage.

Regarding the first stage, when an organization makes an assessment on the opportunity to invest time in a new EU project application, the fact that the entity has a generally strong scientific reputation and/or productivity will not obviously increase its propensity to apply. On the one hand, EU research may be seen as providing opportunities to participate in international academic networks, which will in turn lead to expected benefits in terms of more publications and citations in the future. On the other hand, however, as researchers often point out, participation in EU projects is costly, requiring substantial time for managing and reporting

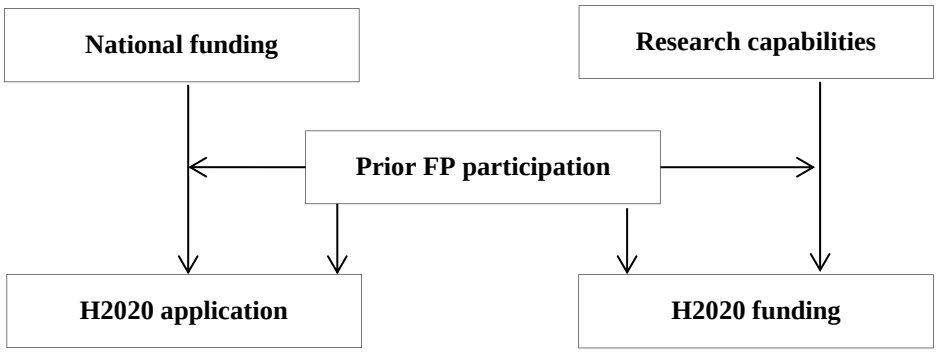
procedures. Anticipating these costs, researchers—and talented and productive ones in particular—may be reluctant to become involved in EU applications because this type of activity may divert time and energy from basic and internally-funded academic research. If so, we argue, strong research capabilities will not increase an organization's propensity to apply—perhaps the converse.

When we shift the focus to the second stage of the participation process, however, the above arguments do not hold. In the selection procedure of determining which applications to accept for funding, research capabilities (reputation and/or productivity) will be a core element that increases the probability of success. Furthermore, it is reasonable to argue that in EU-FP programs and international research in general, well-reputed organizations will be more willing to cooperate with other well-reputed organizations *(homophily*). In EU applications it is the reputation of the entire research consortium that is evaluated, making research capability (and in particular reputation) all the more relevant. We therefore formulate the following hypotheses:

H3.1: *Stronger research capabilities do not increase the probability that an organization will apply for H2020 funding*.

H3.2: *Stronger research capabilities increase the probability that an organization will succeed in getting H2020-funded projects*.

Figure 1 Theory framework and hypotheses



3. Context and data

3.1 The Norwegian research system

Policymakers often compare the national research performance of Norway with that of other small economies in Europe like Denmark, Finland, Sweden, the Netherlands and Austria, as the systems and funding mechanisms are similarly structured, to a certain extent (Norwegian Ministry of Education and Research 2015; Langfeldt et al. 2012).

The Norwegian R&D sector consists of multiple university colleges of varying size, eight large universities, a large number public research organizations, hospitals,[1] state entities and a broad array of private firms—all with the possibility of participating in EU FP (Research Council of Norway 2015). Ever since Norwegian researchers began to become involved in EU FPs, national budgetary appropriations for participation and mobilizing schemes have increased considerably (Langfeldt et al. 2012; Gornitzka and Langfeldt 2008). A declared goal is to increase the degree of R&D funding from EU FPs, at the governmental policy level and the institutional level (Norwegian Ministry of Education and Research 2014; Gornitzka and Langfeldt 2008). In the 7th Framework Program, Norwegian researchers were awarded 1.67% of the available EU funding—whereas the official goal is 2% by the end of Horizon 2020 (Norwegian Ministry of Education and Research 2014). The steadily growing importance of internationalization of research, especially vis-à-vis the EU FP, has received considerable political interest (Langfeldt et al. 2012).

For policymakers, in Norway and elsewhere, there is considerable interest in improving project participation rates and in understanding the factors that obstruct or drive engagement and grant funding in EU FP. To increase participation, the Research Council of Norway (RCN) has established and administers two specific EU-mobilizing incentives. Both aim at getting more Norwegian R&D organizations to apply for EU FP funding. The Project Establishment support scheme, "PES2020," is available to all organizations in Norway that conduct research it is a financial instrument whereby research organizations can literally *buy* time to draft better proposals to EU FP. The second incentive is also a financial mechanism, and has been exclusively developed towards PROs. As EU FP project funding never covers the full costs of a project (given the high cost levels in Norway) this leaves many Norwegian PROs, particularly those with little or no core funding, with project-budget deficits. The "STIM-EU" program is aimed at avoiding this. The scheme provides a research entity with additional funding, based on the amount of total EU FP funding granted. Even though funding is granted only after an EU FP proposal has been accepted, the existence of the STIM-EU scheme incentivizes organizations to apply. In addition, the RCN administers a wide array of programs and schemes structured so as to have an indirect effect on mobilizing for EU FP participation. These are intended to strengthen the capacities of researchers and their organization, ultimately making them stronger competitors for subsequent EU FP grants. Here it should be noted that all of these policy funding schemes are implemented by the Norwegian government in order to encourage research organizations' to participate in EU research, and that they target organizations rather than individual researchers. This is consistent with the approach adopted in this paper, of focusing on the organizational level as the main unit of analysis, and not on the individual researchers or teams.

3.2 Data

Our analysis is based on data on project applications to the 8th European Framework Program Horizon 2020, using data from the European Commission's data warehouse, ECORDA. The database contains information similar to what is publicly available in the Community Research and Development Information Service (CORDIS; cordis.europa.eu), particularly for data on participations in the 6th and 7th FPs. However, for the current FP, Horizon 2020, data in the CORDIS database are not readily available, nor sufficiently updated.

[1] Hospitals without university function are formally a part of the PRO sector, whereas those with university instruction functions fall within the HEI sector (Research Council of Norway 2015).

Access to ECORDA is restricted to the Commission itself and to the Ministries and Research Councils in the member- and associated countries. ECORDA contains applicant information covered by rules of confidentiality, unless presented in aggregated form—particularly for information on applicants that were rejected.

The dataset we have collected for the present study covers 1,402 applications made by Norwegian research organizations for the first period of the framework program (2014–February 2015)[2]. Applications from profit-seeking enterprises such as private companies and state entities (municipalities and other public authorities) have been excluded. Our dataset focuses solely on non-profit research organizations, consisting of higher education institutions (HEIs, 810 applications), PROs (516), and hospitals (76).

Previous studies of EU FP participation have focused on HEIs, with the exception of a few grey papers (Godø et al. 2009). However, for countries like Norway, Germany, Italy and France, where a large portion of the research sector is made up of PROs, the inclusion of these alongside HEIs is warranted, as noted by Geuna (1998). Thus, our study offers an important addition to the literature.

Of the total population of 1,402 applications, 192 are registered as "granted", 1,110 as "rejected," and 100 as "reserved". The "reserved" status means that the EC has placed the application on a waiting list in case the grant winner is removed for some reason. In some (rare) instances, an applicant on the reserved list has been offered the grant. This means that some applications classified as "reserved" might sometimes receive funding, but in our study these are not considered as part of the "successful" group. In the retrieved data, organizations are listed by their legal name. It was a relatively straightforward matter to exclude profit-seeking organizations, like private firms, since we have information of the full population of HEIs and PROs.

The application data offer unique insights on the success and failures of organizations applying for funding. In previous studies (Geuna 1998; Lepori et al. 2015; Nokkala et al. 2011), only information on successful applications has been used for analysis. Recently, Lepori et al. (2015) used the large EUPRO database (Roediger-Schluga and Barber 2008; Primeri and Reale 2012; Protogerou et al. 2010), which contains highly detailed information on applications, but only on those that were granted funding. This does not make it possible to tell whether an organization that has not been granted project funding submitted an application in the first place. That puts non-successful applicants in the same category as those who choose not to apply at all, making it impossible to distinguish between the stage of self-selection (application decision) and the stage of actual EC selection. By contrast, our dataset can identify whether any given Norwegian research organization has applied to the H2020 program, whether it has received a rejection or an approval decision, and how many projects it has been granted. Hence, our database enables a two-stage analysis of the participation process outlined in the previous section.

Data from Horizon 2020 applications were matched with data from several additional sources. Two central national datasets have been used. First, we collected publicly available R&D statistics at the organizational and faculty level for 2013 from the Nordic Institute for Studies in Innovation, Research and Education (NIFU: foustatistikkbanken.nifu.no). NIFU collects data similar to the European Tertiary Education Register (ETER), but with more details on Norwegian R&D organizations. We collected information on the whole population of HEIs (41 university colleges and 8 universities), hospitals, including university hospitals

[2] The data retrieved were entered into ECORDA in July 2015, but it covers project applications only until February 2015. The EU Commission has decided that there must be a time-lag of approx. 5 months from the application deadline until data are published.

(36), and PROs (96) in Norway. Eight of the HEIs are large universities. We decided to treat these at the faculty level, so as not to skew the data, as these universities are large in terms of productivity, funding and staff size, when compared to other HEIs and PROs. All subsequent information for these large universities in the analysis is thus at the faculty level (the number of faculties in Norway is 58).

Second, to assess possible complementarities between national R&D schemes and EU FP, we collected information on all RCN research project grants for the period 2013–2015. We also collected information on the specific support schemes emplaced buy the Norwegian authorities to encourage participation in EU FP, such as STIM-EU and PES2020 (see section 3.1). Further, we could control for the indirect effect of thematic research programs at the RCN that might have complementarity effects with EU-funded research, such as the research program KLIMAFORSK. Other thematic programs might possibly also be relevant, but we did not have sufficient observations to include them in the analysis.

Since the additional data sources noted above provide data at the organizational level (rather than the application level), we have had to aggregate our application dataset at the organizational level in order to combine these different data sources in a single database (as done by other previous studies on EU participation). In total, our dataset for the empirical analysis consists of 231 organizations (99 HEIs, 36 hospitals, and 96 PROs). The advantage of this empirical strategy is that it enables us to examine a rich variety of explanatory factors from these different data sources. A drawback, however, is that by focusing on the organizational level we are unable to study characteristics related to the individual application and/or researchers, and how these interact with other explanatory dimensions in a multi-level setting.

3.3 Variables

The dependent variable in the first step of the analysis is a binary measure indicating *if an organization has applied for a H2020 research project or not.* The dependent variable in the second step of analysis is a count variable indicating *the number of successful applications (i.e. projects granted funding).*

Regarding the main explanatory variables, *FP6 funding* and *FP7 funding* are indicators of prior participation (granted projects) to the 6th FP (2003–2006) and the 7th EU FP (2007–2013).[3] The next two explanatory variables account for organizational research capabilities. *Reputation* is the average number of citations per publication for the period 2010–2014. The bibliometric data were derived from Elsevier's SciVal database (scival.com).[4] *Publication points* are the number of publication points awarded to an institution for the year 2013, divided by the number of full-time equivalent academic staff (FTEs). In Norway, a part of the funding system for research organizations is based on the number of publications that an institution produces each year. These publications are awarded points according to how highly a journal is ranked. Performance of

[3] In additional exercises not reported in this paper, we used two corresponding variables indicating whether an organization has served as coordinator of an FP6 or FP7 project, in order to take into account the role of network centrality in previous participation. These results confirm the positive role of network centrality for participation in EU research. We have also sought to use "coordinated projects" (rather than mere participation) as the dependent variable in the regressions. However, our sample includes only a limited number of coordinated projects, so these additional exercises should be interpreted with caution.

[4] For this variable, we have positive citation data for only 89 institutions, because of the set threshold of minimum 500 publications for inclusion in this database. We set "reputation" for missing organizations at 0, as done by Lepori et al. (2015).

each institution is registered and publicly available online from the Norwegian Scientific Index (CRIStin; cristin.no).

Shifting the focus to the variables measuring national support and other (non-EU) funding means, the main variables that we use in our regression model are two: (1) *national funding*, which includes funds from the Norwegian Research Council and other national funding agencies (excluding core and private company funding); (2) PES2020, STIM-*EU*, and *KLIMAFORSK*, which are dummy variables describing whether the organization has received funding from these specific RCN instruments during the period 2013–2015. As explained, *STIM-EU* and *PES2020* are direct financial schemes aimed at encouraging applications to EU FP. To control for additional funding means, we include two variables measuring the *share of funding received from private companies*, and the *share of funding received from other organizations abroad (excluding firms and EU)*.[5]

Among the set of control variables, we include first of all the *scientific orientation* and *size* of the organization, as both have been shown to be relevant determinants of participation in EU FPs (Lepori et al. 2015; Hakala et al. 2002; Geuna 1998). *Scientific orientation* variables are measured as the amount of funding allocated to the specific scientific discipline to which the organization belongs, as a share of total funding. This is a pre-constructed measure in Norwegian R&D statistics, indicating the dominant scientific specialization of the institution (social sciences, humanities, medicine and health, technology, mathematics and natural sciences). *Size* is the number of full-time equivalents (FTE) of academic staff working in R&D activities. This measure excludes technical and administrative staff as well as time spent on teaching. Data were retrieved from the R&D statistics for 2013.

Next, as prior studies on EU FP determinants have focused mainly on HEIs (e.g. Lepori et al. 2015; Geuna 1998; Nokkala et al. 2011), we wished to see whether HEIs and PROs (both included in our dataset) behave differently. We therefore included a dummy variable to control for organizations classified as HEIs in the application data (1: HEI; 0: PRO). Furthermore, to control for *regional location* of the research organization, we add dummy variables for Norwegian regions in accordance with the Eurostats NUTS 2 classification system (Eurostat 2008).[6]

Table 1 provides descriptive information on the variables included in our analysis, and Table 2 reports the correlation coefficients. Table 1 shows that 56% of Norwegian research organizations applied for participation in Horizon 2020 during the first period of this new EU FP program (between 2014 and February 2015). Further, 130 organizations applied to H2020, and of these, 62 organizations have been granted funding for participation in at least one project (average 0.83 projects).

[5] Lepori et al. (2015) used external funding as an indicator for third-party funds, including external funding from the EU, Research Councils and private companies. We have chosen to treat these variables separately and divide them by total funding to get a measure of the amount of funding by total revenue.

[6] Some of our explanatory variables have skewed distribution, and in particular those measuring R&D capabilities, financial conditions, and size. We log transformed these variables before entering the regression model.

Table 1 Descriptive statistics

	N	Mean	Std.Deviation	Min	Max
Dependent variables					
Application (dummy)	231	0.560	0.497	0	1
Granted projects (count)	149	0.83	2.480	0	25
Persistence					
FP7 participation (count)	231	5.34	16.515	0.00	206
FP6 participation (count)	231	3.02	7.973	0.00	74
R&D capabilities					
Reputation[1]	231	3.045	4.351	0.00	18.4
Publication points[2]	183	1.207	0.904	0.00	6.4
National funding schemes					
National funding (percent)	222	18.329	18.874	0.00	100
PES2020 (dummy)	231	0.280	0.449	0	1
STIM-EU (dummy)	231	0.140	0.351	0	1
Klimaforsk (dummy)	231	0.190	0.397	0	1
Firm funding (percent)	222	9.314	17.643	0.00	91.3
International funding (percent)	222	2.918	9.029	0.00	100
Control variables					
Size[3]	222	91.783	147.458	0.50	1371.8
Higher education institution (dummy)	231	0.430	0.496	0	1
Humanities [4]	231	0.108	0.246	0.00	1
Social sciences[5]	231	0.277	0.340	0.00	1
Mathematics and natural sciences[6]	231	0.145	0.236	0.00	1
Medicine[7]	231	0.162	0.280	0.00	1
Technology[8]	231	0.259	0.372	0.00	1
Regional dummies					
Oslo and Akershus	231	0.424	0.495	0	1
Østlandet	231	0.091	0.288	0	1
Agder and Rogaland	231	0.074	0.262	0	1
Vestlandet	231	0.156	0.364	0	1
Trøndelag	231	0.143	0.351	0	1
North Norway	231	0.113	0.317	0	1

Notes: [1)] average number of citations (2010-14); [2)] number of publication points / Size; [3)] number of full-time equivalents in R&D; [7,5,6,7,8)] amount of funding allocated to a specific scientific discipline by share of total funding.

Table 2 Correlation matrix

	Application	Granted projects	Reputation	Publication points	Firm funding	National funding	International funding	FP7 participation	FP6 participation	STIM-EU	PES2020	Klimaforsk	Size	HEI	Humanities	Social sciences	Math. & nat. sciences	Medicine	Technology
Application	1	0.526***	0.404***	-0.079	0.411***	0.388***	0.432***	0.618***	0.510***	0.310***	0.409***	0.367***	0.613***	0.181***	-0.066	0.150**	0.391***	0.265***	-0.071
Granted projects		1	0.326***	-0.039	0.267***	0.378***	0.381***	0.645***	0.614***	0.378***	0.303***	0.374***	0.525***	0.044	-0.013	0.020	0.392***	0.152**	-0.013
Reputation			1	0.024	0.132**	0.165**	0.271***	0.440***	0.418***	0.004	0.043	0.195***	0.656***	0.508***	0.239***	0.075	0.582***	0.313***	0.060
Publication points				1	-0.403***	-0.109	-0.221***	-0.074	-0.111	-0.259***	-0.092	0.029	-0.164**	0.267***	0.095	0.036	-0.156**	0.039	-0.086
Firm funding					1	0.434***	0.442***	0.348***	0.332***	0.335***	0.414***	0.173***	0.351***	-0.121*	-0.154**	0.070	0.215***	0.141**	-0.076
National funding						1	0.502***	0.461***	0.531***	0.378***	0.398***	0.387***	0.327***	-0.130*	-0.180***	0.330***	0.288***	-0.037	-0.146**
International funding							1	0.532***	0.573***	0.419***	0.440***	0.307***	0.409***	-0.097	-0.124*	0.137**	0.322***	-0.061	-0.038
FP7 participation								1	0.693***	0.484***	0.417***	0.402***	0.657***	0.016	-0.023	0.033	0.428***	0.171***	0.019
FP6 participation									1	0.366***	0.341***	0.488***	0.590***	-0.026	-0.064	0.033	0.381***	0.029	-0.057
STIM-EU										1	0.438***	0.205***	0.205***	-0.329***	-0.228***	-0.008	0.136**	-0.132**	-0.002
PES2020											1	0.257***	0.220***	-0.165**	-0.171***	0.125*	0.121*	0.065	-0.023
Klimaforsk												1	0.323***	-0.006	-0.058	0.077	0.386***	-0.089	-0.076
Size													1	0.343***	0.218***	0.197***	0.533***	0.411***	-0.102
HEI														1	0.488***	0.311***	0.378***	0.329***	-0.009
Humanities															1	0.156**	0.179***	0.202***	-0.164**
Social sciences																1	0.043	0.053	-0.357***
Math. & nat. sciences																	1	0.184***	0.052
Medicine																		1	-0.237***
Technology																			1

* Correlation is significant the 0.10 level, ** at the 0.05 level, or *** at the 0.01 level (2-tailed)

4. Results

4.1 Econometric model

Our empirical model is estimated in two steps. The first stage estimates the probability that an organization applies for an H2020 research grant (dependent variable: APPLICATION dummy). This first step is carried out on the whole population of Norwegian HEIs and PROs. The second stage focuses solely on the sub-sample of research organizations that have applied for H2020 research grants. It estimates the probability that an organization will be granted one or more EU-funded projects (dependent variable: GRANTED, count indicator). The model is then specified by the following two equations:

$$\Pr\{APPLICATION_i\}=\alpha + \beta\ PERSISTENCE_i + \gamma\ RDCAPABILITY_i + \delta\ NATFUNDING_i + \eta[CONTROLS_i]+\varepsilon_i \quad (1)$$

$$Nr\{GRANTED_i\} = \zeta + \theta\ PERSISTENCE_i + \lambda\ RDCAPABILITY_i + \pi\ NATFUNDING_i + \rho[CONTROLS_i]+\mu_i \quad (2)$$

The explanatory variables included in the two equations are the same: previous participation in EU FP (PERSISTENCE), research capabilities, reputation and scientific productivity (RDCAPABILITY), and national funding conditions (NATFUNDING). The vector of control variables (CONTROLS) include, as indicated in the previous section, indicators of organizational size, scientific domain, type of organization (HEI vs. PRO), and regional location. We include the same set of regressors in the two equations because we want to investigate the extent to which the factors highlighted in previous research have different effects on the two stages of the participation process. We also include in the two equations a set of interaction variables in order to analyze whether prior participation in FP moderates the effect of the other variables on the likelihood of applying and/or being awarded an H2020 project.

The advantage of this econometric specification over previous works is twofold. First, equation 1 is meant to provide new insights on the factors that affect research organizations' propensity and motivation to apply for EU-funded projects, in turn offering indications as to what national and European authorities could do in order to mobilize domestic organizations towards greater participation in H2020. To the best of our knowledge, this aspect has not been investigated in previous research, probably due to the unavailability of data that can distinguish applicants from non-applicants.

Second, equation 2 has been analyzed in previous works (e.g. Geuna 1998; Lepori et al. 2015). However, while such research has typically estimated this type of equation on the entire population of research organizations (including applicants as well as non-applicants), we focus solely on the sub-sample of project applicants. This, we hold, can provide an unbiased estimate of the effects of the explanatory variables of interest on the probability of receiving H2020 funding.[7]

[7] In previous research, due the absence of data enabling to distinguish between non-applicants and unsuccessful applicants, both of these groups were combined together and typically given a value of 0 in the regressions, whereas successful applicants were given a positive integer value (1 if the dependent variable was defined as a dummy, and values equal or

We estimate equation 1 using a logit model, and equation 2 using a zero-inflated Poisson regression model (ZIP). The ZIP model is useful when there is a fair amount of null observations in the outcome variable (Greene 1994). As the success rate for project applications to EU FP is quite low, many observations receive the value of 0 (no funded application). The ZIP model accordingly accounts for such a high presence of null observations. The Vuong test in our regressions confirms that the ZIP specification is significantly better than a standard Poisson regression (Perumean-Chaney et al. 2013). The test was significant (P= <0.001) in all regressions.[8]

greater than 1 if the variable was measured as a count). However, combining together non-applicants and unsuccessful applicants in the same regression tends to underestimate the estimated slopes of interest (the larger the mean of a given explanatory variable for the group of non-applicants, the larger the bias). We overcome this problem by estimating equation 2 only on the sub-sample of applicants. Table 3 reports the median of our explanatory variables for four groups of observations: (1) non-applicants, (2) unsuccessful applicants, (3) applicants that have been awarded one project, (4) applicants that have been awarded more than one project. The Table shows that, for the group of non-applicants, three variables (publication points, national funding, and size) have positive median. Hence, our econometric approach yields more precise estimates for these three variables in particular.

[8] We have also estimated our model by following different econometric approaches. First, we have reproduced the approach used by Geuna (1998) and Lepori et al. (2015) on our dataset, i.e. estimating equation 2 by means of a logit (for a dummy dependent variable: funded vs. not-funded), and by means of a truncated regression (for a count-dependent variable that excludes organizations with 0 funded projects). The problem with these approaches is that, due to the relatively small size of our sample, the variability is limited and it is difficult to obtain precise results. Second, we have estimated equations 1 and 2 jointly by means of a Heckman sample selection model (Heckman 1979). However, this approach requires the dependent variable in the second step to be dichotomous. Given the fairly small sample in our study, using a dummy-dependent variable greatly diminishes the data variability. Sample selection models that account for count outcomes are still under development and are not yet incorporated in standard econometric software. We tried using the sample selection model wrapper (SSM) based on generalized linear latent and mixed models (GLLAMM), developed by Miranda and Rabe-Hesketh (2006), but the procedure failed to produce the required output.

Table 3 Median values of main explanatory variables by group (H2020 participation)

Variables	No application	Applied—no success	Granted: 1	Granted: 2 or more
FP7 participation	0,000	1,000	4,500	15,000
FP6 participation	0,000	0,000	2,000	8,500
Reputation	0,000	1,450	2,750	6,350
Publication points	1,231	0,972	1,034	0,951
National funding	5,855	12,247	19,317	27,790
PES2020	0,000	0,000	0,000	1,000
STIM-EU	0,000	0,000	0.000	0,500
Klimaforsk	0,000	0,000	0,000	1,000
Firm funding	0,000	2,440	5,327	5,679
International funding	0,000	0,287	1,274	1,486
Size	17,400	56,400	85,100	169,770
Higher education institution	0,000	1,000	0,500	0,000
N	**101**	**68**	**26**	**36**

Table 4 Comparison of mean values between the groups of successful and unsuccessful applicants (group variable: 1= at least one project funded, otherwise 0).

	Mean (Std. deviation)			*t-test*		
Variable	**Successful** (n=62)	**Not successful** (n=68)	Mean difference	t	Df	Sig. (2-tailed)
Granted project	3.10 (4.00)	0 (0)	3.097	6.085	61.000	<0.001
FP7 participation	16.92 (28.76)	2.35 (3.21)	14.566	3.966	62.389	<0.001
FP6 participation	9.29 (13.26)	1.44 (2.60)	7.849	4.579	65.273	<0.001
Reputation	5.13 (4.69)	3.74 (4.28)	1.385	1.759	128	0.081
Publication points	1.07 (0.55)	1.08 (0.65)	-0.014	-0.122	113	0.903
National funding	26.27 (15.95)	19.10 (18.58)	7.167	2.309	123	0.023
PES2020	0.48 (0.50)	0.40 (0.93)	0.087	0.992	128	0.323
STIM-EU	0.35 (0.48)	0 (0)	0.034	3.056	128	0.003
Klimaforsk	0.42 (0.49)	0.24 (0.42)	0.184	2.252	120.876	0.026
Firm funding	11.42 (16.12)	10.65 (18.72)	0.766	0.245	123	0.807
International funding	3.87 (6.01)	2.34 (5.06)	1.530	1.542	123	0.126
HEI	0.50 (0.50)	0.60 (0.49)	-0.103	-1.177	128	0.242
Size	200.63 (225.62)	86.88 (91.87)	113.739	3.659	78.594	<0.001
Humanities	0.04 (0.06)	0.07 (0.18)	-0.028	-1.208	87.269	0.230
Social sciences	0.26 (0.29)	0.32 (0.34)	-0.065	-1.156	128	0.250
Mathematics and natural sciences	0.27 (0.29)	0.16 (0.26)	0.101	2.126	128	0.035
Medicine	0.17 (0.23)	0.22 (0.34)	-0.052	-1.030	117.772	0.305
Technology	0.17 (0.24)	0.17 (0.25)	-0.004	-0.085	128	0.932
Oslo and Akershus	0.45 (0.50)	0.46 (0.50)	-0.004	-0.048	128	0.961
Østlandet	0.02 (0.13)	0.15 (0.36)	-0.131	-2.835	85.109	0.006
Agder and Rogaland	0.05 (0.22)	0.06 (0.24)	-0.010	-0.261	128	0.794
Vestlandet	0.19 (0.39)	0.08 (0.29)	0.105	1.718	109.691	0.089
Trøndelag	0.19 (0.39	0.13 (0.34)	0.061	0.943	128	0.348
North Norway	0.09 (0.29)	0.12 (0.32)	-0.02087	-0.381	128	0.704

Notes: For a definition of these indicators see table 1 above.

4.2 Estimation results

Tables 5 and 6 present the results of the econometric estimations of equations 1 and 2, respectively. Before discussing the results of the tests of the hypotheses formulated in section 2.2, let us take a brief look at the estimated coefficients of some important control variables. In particular, SIZE has often been pointed out as a relevant explanatory factor in previous research (Lepori et al. 2015). Our estimates show an interesting difference in the effects of this variable in the first and in the second equation. The variable has the expected positive sign in equation 1 (see Table 5), and the estimated coefficients are large and significant in all regressions. By contrast, the variable has lower and weakly significant estimated coefficients in equation 2 (see Table 6).[9] We find a similar pattern for the HEI dummy variable that distinguishes higher education institutions from PROs. Taken together, these results indicate that larger research organizations, and HEIs in particular, are more likely to apply for EU FPs, arguably due to the economies of scale they can enjoy in the phases of developing, managing and administering new applications to H2020. These may lead to cost advantages that can prove crucial in the first stage of the participation process (self-selection), but not necessarily in the subsequent selection process (where it is quality, more than size, that matters).

Turning now to our set of main explanatory variables, the indicator of prior participation in EU FPs (PERSISTENCE) has positive and estimated coefficients in both Table 5 and Table 6.[10] The magnitude of these estimated coefficients is one of the largest in the battery of regressors, confirming that prior participation in EU FP is a major factor that strengthens both the propensity of Norwegian research organizations to apply to H2020 *and* their ability to get H2020 projects. As note in our propositions H1.1 and H1.2 (see section 2.2), the variable is highly relevant in both stages of the participation process—although, in our view, for different reasons. In the first (self-selection) stage, previous participation in an EU-funded project means that the organization will not have to sustain the sunk costs of amassing information and knowledge on the application procedure, as well as network and team building costs. In the second stage, prior participation matters because an organization that has previously been involved in an EU project can be assumed to have benefited from learning effects regarding the application procedure, EU research priorities, and the quality of potential partners—and these learning effects will improve the quality of the application and its chances of success.

The second set of explanatory variables of interest relates to national funding characteristics. The specific hypothesis formulated in section 2 is that that the structure of national funding matters in the first stage of the participation process, but not equally so in the second. To test these propositions (H2.1 and H2.2), let us compare estimation results for the relevant variables in Table 5 and Table 6. The national funding variable (share of research grants received from national funding agencies) does not prove to have a significant effect in equation 1. Interestingly, however, the variable is important and significant when we interact it with the

[9] A possible reason why the size variable is not significant in the estimations in step 2 may be due to multicollinearity. The correlation coefficients in table 2 indicate SIZE to be positively correlated with the FP6 and FP7 participation variables (although the VIF statistics for these variables are below critical threshold levels for multicollinearity). If we exclude the latter from the regressions, the size indicator becomes statistically significant. To test further the effect of size on H2020 participation, we have also carried out another exercise based on a matching approach. The results of matching results, reported at the end of this section, show indeed that size has a significant correlation to H2020 participation.

[10] A cross-classification table between participation to FP7 and H2020 shows that persistence is important in our sample, as expected, although it does not completely predict participation patterns. In fact, 24% of organizations that got at least one project in FP7 also got funding from H2020, whereas 25% were not able to succeed in H2020. On the other hand, 48% of the organizations that did not have any funded project in FP7 did not have any H2020 project either; and only 3% of organizations managed to get a H2020 project although they did not have any FP7 project.

PERSISTENCE variable (prior participation in EU FP7). This pattern indicates that previous participation positively moderates the effect of national funding availability on the propensity to submit an application for the new program H2020. In other words, the expected complementarity effects between nationally-funded and EU-funded research emerge more clearly for organizations that have previously sustained the costs of sunk information and network building (in other words: if an organization already knows how to play the EU game, having complementary national funds available will further encourage it to continue to participate). Table 5 also confirms that the variable *PES2020*, one of the specific policy schemes in place in Norway to encourage participation in EU-funded research, has positive, significant and strong estimated coefficients.

We note the corresponding results for equation 2 (Table 6): the national fund availability variable is not significant (neither alone nor in interaction with PERSISTENCE), whereas the *PES2020* variable is still positive, significant, but its estimated coefficient has a much lower magnitude than the corresponding results in table 5. On the whole, these patterns are in line with the hypotheses formulated in section 2.2, that external funding granted by national institutions increases the probability of an organization submitting an H2020 application (by making available additional resources to invest in the application process), but does not substantially increase the probability that an organization will succeed with its H2020 application (since that outcome depends more on the quality of the application than on the financial conditions of the applying entity).

Finally, we shift the focus to the results on the variables accounting for organizational research capabilities: academic reputation (measured by citations and impact factor) and scientific productivity (measured by the number of publication points in the recent past). In section 2 we argued that research capabilities matter differently in the two stages of the participation process, making it therefore important to distinguish these impacts: we hold that this dimension will be relevant only in the second (selection) stage (in line with the results of previous studies), and not in the first stage. To test this, let us compare again the estimation results for equation 1 and equation 2. In Table 5, the reputation variable emerges as not significantly correlated to the propensity to apply for EU funding. The scientific productivity (publication points) variable is not significant either. Interestingly, however, when we interact this indicator with the PERSISTENCE variable, we find a negative, strong and significant effect. One interpretation could be that researchers often perceive involvement in EU projects as a costly, demanding, and bureaucratic process, and may therefore be reluctant to participate. To the extent that more productive researchers concentrate on academic recognition and publication activity, high scientific productivity will not make an organization more likely to apply to H2020—maybe even the converse.

However, as expected, the results of the estimations for equation 2 tell a different story, and one more in line with the results of previous research. Table 5 shows that the reputation variable, when interacted with the indicator of prior participation in FP6, has a positive and significant estimated coefficient. This suggests that scientific reputation does matter in the selection of proposal submitted to H2020, especially if an excellent reputation is combined with learning effects arising from having previously participated in EU-funded projects. Although the reputation variable emerges as important in equation 2, the publication variable is not significant, which would indicate that having a large number of publications does not increase the likelihood of success in the H2020 contest, whereas having a high scientific impact (reputation) does. In summary, in line with our theory arguments (H3.1 and H3.2), we find that stronger research capabilities (as measured by reputation) do not increase the probability that an organization will submit an H2020 application, but they do increase the

probability of success in getting H2020-funded projects.

A possible issue in our regressions is that some of the variables of interest – and in particular reputation, size and FP7 participation are positively correlated with each other. In our relatively small sample, this may make it hard to disentangle the effect of each of these variables on H2020 participation. The resulting multicollinearity patterns can lead to concerns that the effect of these variables cannot be estimated with statistical precision in a multivariate regression model given the relatively small size of our sample. We have therefore carried out an additional exercise in the attempt to disentangle the effects of the aforementioned variables, focusing on step 2 of the econometric model.

In Table 7, we report the results of a propensity score matching analysis that we have carried out in order to provide an alternative and more precise estimation of the effect of these variables. The main idea of the matching approach is to consider the explanatory variable of interest as a "treatment indicator", and estimate the effect of the treatment variable on the outcome variable (H2020 participation) (Caliendo and Kopeinig 2008; Cerulli 2010). More precisely, the propensity score matching exercise followed these steps: (1) We focused on one explanatory variable at a time and considered it as a treatment variable. (2) We transformed this variable into a dummy indicator (with value 1 for observations above the median, and 0 for observations below the median). (3) We estimated the predicted probability from a probit regression of the treatment dummy on a set of covariates (i.e. the same set of covariates used for the estimation of equation 2); this predicted probability is the so-called *propensity score*. (4) The matching algorithm (carried out in the statistical software Stata) then used the propensity score to construct two groups of observations: a treated and a control group. The identification of these two groups was based on the nearest-neighbour method, and the Mahalanobis distance metric. (5) The average treatment effect on the treated (ATT) was then estimated, which is the difference between the average of the outcome variable (H2020 participation) in the two groups. (6) We then repeated these five steps for the other explanatory variables in equation 2 that are possibly affected by multicollinearity.

The results show that the ATT of FP7 participation on the H2020 dependent variable is positive and significant, similarly to what previously found in our regression results reported in table 6. The publication points indicator does not have a significant ATT effect, a finding that is also in line with our regression results. Interestingly, however, both the reputation and the size variables (in spite of their collinearity with the FP7 indicator) have a positive and signicant estimated ATT effect on H2020 participation. This is a pattern that we were not able to estimate with statistical precision in the multivariate regression analysis, and that the matching procedure reported here makes it possible to identify. These findings are also in line with the econometric results of previous studies in this field on the role of size and reputation for EU project participation (e.g. Lepori et al. 2015). On the whole, these matching results should be interpreted with some caution, due the relatively small size of our sample. However, they provide interesting evidence that extends and corroborates the findings of the multivariate regressions previously presented.

Table 5 Results of step 1. Dependent variable: application (dummy). Estimation method: logit

	Baseline	Interaction 1	Interaction 2	Interaction 3	Full model
FP7 participation	1.302 (0.032)**	-0.788 (0.449)	4.001 (0.012)**	1.587 (0.061)*	2.099 (0.228)
FP6 participation	0.465 (0.481)	0.585 (0.421)	0.375 (0.581)	0.502 (0.450)	0.059 (0.941)
Reputation	0.080 (0.859)	0.266 (0.570)	0.220 (0.639)	0.241 (0.660)	1.062 (0.125)
Reputation * FP7 particip.				-0.275 (0.600)	-0.211 (0.743)
Publication points	0.545 (0.600)	0.604 (0.581)	1.371 (0.222)	0.608 (0.561)	2.608 (0.053)*
Publication points * FP7 particip.			-3.206 (0.039)**		-8.756 (0.002)***
National funding	0.114 (0.751)	-0.126 (0.735)	0.124 (0.746)	0.104 (0.772)	-0.082 (0.848)
National funding * FP7 particip.		0.893 (0.032)**			2.646 (0.002)***
PES2020	1.708 (0.031)**	1.714 (0.041)**	1.771 (0.033)**	1.629 (0.043)**	1.661 (0.077)*
STIM-EU	0.421 (0.717)	-0.482 (0.707)	0.191 (0.881)	0.314 (0.790)	-3.088 (0.129)
Klimaforsk	1.654 (0.114)	1.619 (0.154)	1.575 (0.141)	1.621 (0.123)	1.415 (0.264)
Firm funding	0.805 (0.028)**	0.902 (0.023)**	0.841 (0.034)**	0.872 (0.027)**	1.295 (0.013)**
International funding	-0.586 (0.216)	-0.631 (0.203)	-0.783 (0.128)	-0.605 (0.208)	-1.377 (0.052)*
Size	1.175 (0.015)**	1.365 (0.007)***	1.223 (0.014)**	1.158 (0.016)**	1.863 (0.003)***
Higher education institution	1.826 (0.049)**	1.716 (0.066)*	1.678 (0.073)*	1.861 (0.047)**	1.635 (0.105)
Humanities	-3.269 (0.347)	-3.821 (0.299)	-3.594 (0.348)	-3.286 (0.350)	-6.749 (0.193)
Social sciences	-1.648 (0.607)	-1.623 (0.627)	-1.561 (0.665)	-1.559 (0.631)	-2.084 (0.643)
Mathematics and natural science	-3.426 (0.338)	-4.735 (0.214)	-4.183 (0.302)	-3.634 (0.321)	-9.340 (0.090)*
Medicine	0.240 (0.942)	0.155 (0.964)	0.001 (1.000)	0.277 (0.934)	-0.661 (0.885)
Technology	-2.235 (0.498)	-2.321 (0.502)	-2.812 (0.446)	-2.235 (0.500)	-4.475 (0.339)
Oslo and Akershus	0.481 (0.626)	0.535 (0.592)	0.563 (0.575)	0.428 (0.669)	1.021 (0.363)
Agder and Rogaland	-1.765 (0.165)	-2.117 (0.109)	-1.915 (0.143)	-1.842 (0.152)	-2.747 (0.067)*
Vestlandet	-1.471 (0.179)	-1.671 (0.149)	-1.564 (0.159)	-1.624 (0.157)	-2.324 (0.078)*
Trøndelag	-0.822 (0.452)	-0.818 (0.464)	-1.020 (0.367)	-0.812 (0.459)	-1.356 (0.270)
North Norway	-0.015 (0.990)	-0.110 (0.929)	0.205 (0.866)	-0.063 (0.958)	0.025 (0.985)
-2 Log likelihood	92.463	88.229	88.167	92.203	74.632
Nagelkerke R-squared	0.75	0.76	0.76	0.75	0.81

Observations: 177. P-values in parentheses. Significant at * 0.10 level; ** 0.05 level; *** 0.01 level. A constant is included in all regressions.

Table 6 Results of step 2. Dependent variable: number of granted projects (count). Estimation method: zero-inflated Poisson

	Baseline	Interaction 1	Interaction 2	Interaction 3	Full model
FP7 participation	0.313 (0.068)*	0.253 (0.141)	0.376 (0.027)**	0.297 (0.084)*	0.281 (0.104)
FP6 participation	0.388 (0.032)**	-0.256 (0.582)	0.107 (0.712)	0.060 (0.801)	-0.130 (0.809)
Reputation	0.089 (0.573)	0.125 (0.422)	0.182 (0.260)	-0.282 (0.250)	-0.316 (0.262)
Reputation * FP6 particip.				0.233 (0.023)**	0.242 (0.033)**
Publication points	-0.382 (0.598)	-0.506 (0.477)	-1.554 (0.115)	-0.673 (0.353)	-0.814 (0.401)
Publication points * FP6 particip.			0.364 (0.304)		-0.078 (0.836)
National funding	-0.116 (0.546)	-0.184 (0.410)	-0.234 (0.309)	-0.080 (0.679)	-0.171 (0.497)
National funding * FP6 particip.		0.204 (0.110)			0.072 (0.585)
PES2020	0.778 (0.009)***	0.789 (0.007)***	0.457 (0.114)	0.684 (0.024)**	0.521 (0.081)*
STIM-EU	-0.467 (0.148)	-0.449 (0.160)	-0.234 (0.481)	-0.262 (0.444)	-0.193 (0.575)
Klimaforsk	-0.064 (0.769)	0.035 (0.870)	0.010 (0.962)	0.053 (0.809)	0.109 (0.628)
Firm funding	-0.030 (0.840)	-0.041 (0.777)	0.103 (0.472)	-0.087 (0.575)	0.064 (0.679)
International funding	0.119 (0.559)	0.286 (0.178)	-0.013 (0.943)	0.046 (0.818)	-0.029 (0.878)
Size	0.109 (0.650)	0.075 (0.757)	0.069 (0.787)	-0.080 (0.742)	-0.090 (0.718)
Higher education institution	0.900 (0.090)*	1.236 (0.015)**	0.462 (0.396)	0.907 (0.077)*	0.626 (0.265)
Humanities	-3.109 (0.300)	-4.640 (0.130)	-0.608 (0.840)	-2.053 (0.449)	0.233 (0.938)
Social sciences	0.334 (0.628)	0.110 (0.872)	1.176 (0.087)*	0.345 (0.618)	0.695 (0.327)
Math. & natural science	0.600 (0.307)	-0.075 (0.900)	0.868 (0.175)	0.559 (0.339)	0.409 (0.545)
Medicine	0.740 (0.301)	0.552 (0.444)	0.937 (0.212)	0.829 (0.257)	1.085 (0.153)
Technology	1.110 (0.157)	0.494 (0.502)	1.185 (0.134)	1.288 (0.091)*	0.761 (0.340)
Oslo and Akershus	1.988 (0.069)*	1.316 (0.238)	1.104 (0.341)	1.883 (0.083)*	0.889 (0.440)
Agder and Rogaland	3.152 (0.011)**	2.598 (0.034)**	2.141 (0.090)*	3.098 (0.011)**	1.932 (0.126)
Vestlandet	2.545 (0.022)**	1.873 (0.098)*	1.716 (0.139)	2.385 (0.031)**	1.309 (0.257)
Trøndelag	2.154 (0.053)*	1.629 (0.146)	1.311 (0.259)	2.300 (0.038)**	1.390 (0.230)
North Norway	1.920 (0.094)*	1.091 (0.343)	0.889 (0.446)	1.567 (0.165)	0.424 (0.717)
Log likelihood	-102.652	-101.219	-101.181	-98.624	-95.929
Vuong test of ZIP vs. standard Poisson	5.52 (<0.001)***	5.36 (<0.001)***	5.34 (<0.001)***	6.52 (<0.001)***	6.58 (<0.001)***

Observations: 113. P-values in parentheses. Significant at * 0.10 level; ** 0.05 level; *** 0.01 level. A constant is included in all regressions.

Table 7. Propensity score matching (nearest-neighbour method). Distance metric: Mahalanobis.

Variable	ATT	z	Number treated	Number control	N
Size	1.862	3.17***	58	55	113
Reputation	1.857	3.13***	56	57	113
Publication points	-0.175	-0.38	57	56	113
FP7 participation[1]	1.402	3.35***	87	26	113

Significant at: *** 0.01 level. **Note:** [1)] Dummy. 1 equals to at least one FP7 project participation, 0 otherwise.

5. Conclusions

This paper has analyzed Norwegian research organizations in terms of applying for participation in the EU's Horizon 2020 program during the first months of its implementation. Our dataset consisted of 1,402 applications for the period 2014–2015, both rejected applications and successful ones, matched with R&D statistics for the whole population of higher education institutions and public research organizations in Norway. We analyzed this database by means of a two-step econometric analysis of the likelihood of applying (stage 1) and probability of success (stage 2), focusing on three main explanatory dimensions: prior participation in EU FPs, complementarity with national funding schemes, and scientific capabilities. What have we learned from this exercise—what are the main contributions and implications for research policy?

The main point is that it is essential to distinguish two separate stages in the participation process: the self-selection stage in which a research organization decides whether to invest time and resources in developing a project application, and the selection stage in which the European Commission evaluates and selects proposals. Both conceptually and empirically, these two stages differ, and different factors explain the likelihood of applying and the probability of succeeding in the competition for H2020 funding.

Regarding the first stage, our results indicate two key dimensions that determine whether a research organization is likely to prepare and submit an application: (1) if the organization has previously participated in EU FPs (and thereby already sustained the application-related sunk costs); and (2) the availability of national funding, which may act as a complementary channel and provide additional resources that encourage application efforts. These results provide new scholarly insights on participation in EU-funded research, with important potential implications for policy.

A first implication concerns the national authorities: given the important complementarities between national and international funding, the most effective way for domestic funding agencies to promote internationalization of research is not by outsourcing research funding tasks totally to EU authorities, but by strengthening their own national funding programs. A second implication concerns the EU authorities in charge of the H2020 program. From our results, it appears that this new program has not been able yet to mobilize Norwegian

researchers with strong scientific reputations and high productivity: indeed, we find the likelihood of submitting an application to be negatively related to the organization's scientific productivity. This might have to do with the fact that researchers often see the H2020 program as basically applied in nature and quite demanding in terms of networking and management procedures. Perhaps the EU authorities should consider whether the program should have a more academic focus and simpler procedures, in order to attract a greater number of productive researchers.

Regarding the second stage of the participation process, the main advantage of our econometric approach *vis-à-vis* previous research is that carrying out estimations only on the sub-sample of applicants enabled us to obtain more precise, unbiased, estimates of the relationships of interest here. Our results indicate two main factors that strengthen the likelihood that a research organization will submit a successful application: prior participation in EU FPs (indicating persistence and learning effects from previous EU projects), and scientific reputation. Unlike earlier studies, however, we do not find any significant correlation between publication results and probability of success.

These results give rise to another question for research policy. For national authorities, it is important to emphasize the distinction between the role of academic reputation and the role of scientific productivity. Also in many European countries, including Norway, allocation of national funding to HEIs and PROs is currently based on scientific productivity (number of publication points), among other criteria. However, if the national authorities want to promote the internationalization of their research system and domestic participation in H2020 programs, it would seem advisable to adjust the funding allocation scheme so as to give more weight to academic reputation (quality, impact, citation indexes) and less to the number of publication points as such. This would give a clear signal to research organizations that scientific quality and impact matter more than productivity, spurring further efforts in this direction, and, one hopes, a higher success rate in the H2020 contest.

Finally, we must acknowledge two limitations of this work. The first is that, in seeking to provide a timely analysis of the first available data on H2020 participation, we have not been able to incorporate data for a lengthy period or for the entire duration of the program. This means that our assessment will need to be verified and extended when new data on H2020 participation become available. A second caveat: our analysis has focused on research organizations in Norway, for which rich and high-quality data and information are available. Some of our results might be affected by the specific characteristics of the Norwegian research system, and hence not necessarily hold for a larger sample of other European countries. It is therefore important for future research to reassess our two-step approach and hypotheses, analyzing their validity for the case of other European economies.

References

Annerberg, R., Begg, I., Acherson, H., Borrás, S., Hallen, A., Maimets, T., et al. (2010). Interim Evaluation of the Seventh Framework Programme: Report of the Expert Group. European Commission.

Breschi, S., Cassi, L., Malerba, F., & Vonortas, N. S. (2009). Networked research: European policy intervention in ICTs. *Technology Analysis & Strategic Management, 21*(7), 833-857.

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*(8), 747-772.

Caliendo, M., & Kopeinig, S. (2008). Some practical guidance for the implementation of Propensity score matching. *Journal of economic surveys, 22*(1), 31-72, doi:10.1111/j.1467-6419.2007.00527.x.

Cerulli, G. (2010). Modelling and Measuring the Effect of Public Subsidies on Business R&D: A Critical Review of the Econometric Literature. *Economic Record, 86*(274), 421-449.

Defazio, D., Lockett, A., & Wright, M. (2009). Funding incentives, collaborative dynamics and scientific productivity: Evidence from the EU framework program. *Research Policy, 38*(2), 293-305.

Di Cagno, D., Fabrizi, A., & Meliciani, V. (2014). The impact of participation in European joint research projects on knowledge creation and economic growth. *The Journal of Technology Transfer, 39*(6), 836-858.

Dinges, M., & Lepori, B. (2006). *Public Project Funding of Research Activities: National Differences and Implications for the Creation of a European Research Council.* Paper presented at the Platform fteval Conference–New Frontiers in Evaluation, 24th-25th April,

Eurostat (2008). Statistical Regions for the EFTA countries and Candidate countries. European Commission.

Geuna, A. (1996). The participation of higher education institutions in European Union Framework Programmes. *Science and Public Policy, 23*(5), 287-296.

Geuna, A. (1998). Determinants of university participation in EU-funded R&D cooperative projects. *Research Policy, 26*(6), 677-687.

Godø, H., Langfeldt, L., & Kaloudis, A. (2009). *In Need of a Better Framework for Success: An evaluation of the Norwegian participation in the EU 6th Framework Programme (2003–2006) and the first part of the EU 7th Framework Programme (2007–2008).* Oslo: NIFU STEP.

Gornitzka, Å., & Langfeldt, L. (2008). The internationalisation of national knowledge policies. In A. Gornitzka, & L. Langfeldt (Eds.), *Borderless knowledge. Understanding the "New" Internationalisation of Research and Higher Education in Norway* (Vol. 22, pp. 141-169). Amsterdam: Springer.

Greene, W. H. (1994). Accounting for Excess Zeros and Sample Selection in Poisson and Negative Binomial Regression Models. *New York University, Stern School of Business, Working paper 94-10, 3*.

Hakala, J., Kutinlahti, P., & Kaukonen, E. (2002). Becoming international, becoming European: EU research collaboration at Finnish universities. *Innovation: The European Journal of Social Science Research, 15*(4), 357-379.

Heckman, J. J. (1979). Sample selection bias as a specification error. *Econometrica: Journal of the econometric society*, 153-161.

Henriques, L., Schoen, A., & Pontikakis, D. (2009). Europe's top research universities in FP6: scope and drivers of participation. *JRC Technical Notes* (Vol. 53681): European Commission.

Hoekman, J., Scherngell, T., Frenken, K., & Tijssen, R. (2012). Acquisition of European research funds and its effect on international scientific collaboration. *Journal of Economic Geography*, lbs011.

Langfeldt, L., Godø, H., Gornitzka, Å., & Kaloudis, A. (2012). Integration modes in EU research: Centrifugality versus coordination of national research policies. *Science and Public Policy, 39*(1), 88-98.

Laredo, P. (1998). The networks promoted by the framework programme and the questions they raise about its formulation and implementation. *Research Policy, 27*(6), 589-598.

Lepori, B., Veglio, V., Heller-Schuh, B., Scherngell, T., & Barber, M. (2015). Participations to European Framework Programs of higher education institutions and their association with organizational characteristics. *Scientometrics*(3), 2149-2178, doi:10.1007/s11192-015-1768-2.

Luukkonen, T. (1998). The difficulties in assessing the impact of EU framework programmes. *Research Policy, 27*(6), 599-610.

Luukkonen, T. (2000). Additionality of EU framework programmes. *Research Policy, 29*(6), 711-724.

Luukkonen, T., & Nedeva, M. (2010). Towards understanding integration in research and research policy. *Research Policy, 39*(5), 674-686.

Makkonen, T., & Mitze, T. (2016). Scientific collaboration between 'old'and 'new'member states: Did joining the European Union make a difference? *Scientometrics*, 1-23.

Matt, M., Robin, S., & Wolff, S. (2012). The influence of public programs on inter-firm R&D collaboration strategies: project-level evidence from EU FP5 and FP6. *The Journal of Technology Transfer, 37*(6), 885-916.

Miranda, A., & Rabe-Hesketh, S. (2006). Maximum likelihood estimation of endogenous switching and sample selection models for binary, ordinal, and count variables. *Stata Journal, 6*(3), 285-308.

Nedeva, M. (2013). Between the global and the national: Organising European science. *Research Policy, 42*(1), 220-230.

Nokkala, T., Heller-Schuh, B., & Paier, M. (2011). Ranking lists and European Framework Programmes: Does university status matter for performance in Framework Programmes? In P. T. D. Dill (Ed.), *Public Vices, Private Virtues? Assessing the Effects of Marketization in Higher education* (pp. 111-139). Rotterdam: Sense Publishers.

Norwegian Ministry of Education and Research (2014). Strategy for research and innovation cooperation with the EU - Horizon 2020 and ERA. Oslo.

Norwegian Ministry of Education and Research (2015). Forskningsbarometeret 2015. Oslo.

Okubo, Y., & Zitt, M. (2004). Searching for research integration across Europe: a closer look at international and inter-regional collaboration in France. *Science and Public Policy, 31*(3), 213-226.

Ortega, J. L., & Aguillo, I. F. (2010). Describing national science and technology systems through a multivariate approach: country participation in the 6th Framework Programmes. *Scientometrics, 84*(2), 321-330.

Paier, M., & Scherngell, T. (2011). Determinants of collaboration in European R&D networks: empirical evidence from a discrete choice model. *Industry and Innovation, 18*(1), 89-104.

Perumean-Chaney, S. E., Morgan, C., McDowall, D., & Aban, I. (2013). Zero-inflated and overdispersed: what's one to do? *Journal of Statistical Computation and Simulation, 83*(9), 1671-1683, doi:10.1080/00949655.2012.668550.

Polt, W., & Streicher, G. (2005). Trying to capture additionality in Framework Programme 5—main findings. *Science and Public Policy, 32*(5), 367-373.

Primeri, E., & Reale, E. (2012). How Europe shapes academic research: Insights from participation in European union framework programmes. *European Journal of Education, 47*(1), 104-121.

Protogerou, A., Caloghirou, Y., & Siokas, E. (2010). Policy-driven collaborative research networks in Europe. *Economics of Innovation and New Technology, 19*(4), 349-372.

Research Council of Norway (2015). *Det norske forsknings- og innovasjonssystemet - statistikk og indikatorer*. Oslo: Research Council of Norway.

Roediger-Schluga, T., & Barber, M. J. (2008). R&D collaboration networks in the European Framework Programmes: Data processing, network construction and selected results. *International Journal of Foresight and Innovation Policy, 4*(3-4), 321-347.

Tijssen, R. J. (2008). Are we moving towards an integrated European Research Area? Some macro-level bibliometric perspectives. *Collnet Journal of Scientometrics and Information Management, 2*(1), 19-25.

Ukrainski, K., Masso, J., & Kanep, H. (2014). Cooperation patterns in science within Europe: the standpoint of small countries. *Scientometrics, 99*(3), 845-863.

Viner, N., Powell, P., & Green, R. (2004). Institutionalized biases in the award of research grants: a preliminary analysis revisiting the principle of accumulative advantage. *Research Policy, 33*(3), 443-454.