The important thing is not to win, it is to take part:

What if scientists benefit from participating in research grant competitions?

Charles Ayoubi
Chair in Economics and Management of Innovation - École Polytechnique Fédérale de Lausanne
charles.ayoubi@epfl.ch

Michele Pezzoni
Université Côte d'Azur, CNRS, GREDEG, France;
ICRIOS, Bocconi University, Milan, Italy;
BRICK, Collegio Carlo Alberto, Torino, Italy
michele.pezzoni@unice.fr

Fabiana Visentin*
Chair in Economics and Management of Innovation - École Polytechnique Fédérale de Lausanne ;
BRICK, Collegio Carlo Alberto, Torino, Italy;
fabiana.visentin@epfl.ch

Abstract

"The important thing is not to win, it is to take part," this famous saying by Pierre de Coubertin asserts that the value athletes draw from Olympic games lies in their participation in the event and not in the gold they collect during it. We find similar evidence for scientists involved in grant competitions. Relying on unique data from a Swiss funding program, we find that scientists taking part in a research grant competition boost their number of publications and average impact factor while extending their knowledge base and their collaboration network regardless of the result of the competition. Receiving the funds increases the probability of co-authoring with co-applicants but has no additional impact on the individual productivity.

Keywords: competitive grants, public funding evaluation, scientific productivity, scientific collaboration, learning

JEL codes: I23, O38

***Corresponding author:** Fabiana Visentin, Chair in Economics and Management of Innovation - École Polytechnique Fédérale de Lausanne, Bâtiment ODY, Station 5, CH – 1015 Lausanne. Email: fabiana.visentin@epfl.ch

1. Introduction

Throughout the years of economic history, the benefits of competition in terms of social welfare and knowledge production have been debated (Arrow, 1962; Aghion et al. 2005). In contexts where only part of the competitors gets all the monetary reward, competition takes on the characteristics of a race. Research grant competitions offer a stylized example: the scientists who submit the most convincing proposal to the funding agency 'win'. However, is this competition creating a winners-take-all situation? Or do participants find any benefit in only taking part in the race? Scientific grants are a convenient setting to evaluate the benefits and drawbacks of competition for all contestants, winning or not. We use unique data on a Swiss grant and find that merely taking part in a competition is useful regardless of the result. Specifically, this paper is the first to bring empirical evidence on an overlooked aspect of the research grant process, i.e. the effect of taking part in a grant competition on the scientific productivity, learning, and collaboration of scientists. Furthermore, we complement the extant literature on the impact of receiving funds from a public funding agency.

When conducting research work, scientists are guided by financial remuneration, puzzle solving satisfaction and search for fame and glory (Stephan, 1996). However, nowadays, regardless of their initial motivation, scientists need substantial funding to produce science. Hence, the ability to raise funds is becoming a key skill in managing research laboratories (Etzkowitz, 2003) and a base in the evaluation of scientists' performances along with publication records (Ruben, 2017). Researchers spend an increasing number of hours in writing grant proposals with an uncertain outcome, and, when awarded, in managing the resources they receive. Developing empirical evidence on the benefits and drawbacks of these time-consuming activities would support policy makers and funding agencies in crafting funding systems. However, to our

knowledge, extant literature does not include any analysis of the impact of the application process, and studies evaluating funding efficiency are still scarce.

The debate in the scientific community on the opportunity to participate in research grant competitions mainly focuses on the costs they entail. Ioannidis provocatively stated that "the research funding system is broken: researchers don't have time for science anymore. Because they are judged on the amount of money they bring to their institutions, writing, reviewing and administering grants absorb their efforts" (Ioannidis, 2011). Similarly, Stephan (2010) claimed that "grant applications divert scientists from spending time doing science" and reported an insightful example: "a funded chemist in the U.S. can easily spend 300 hours per year writing proposals". She added that "while some of this effort undoubtedly generates knowledge, much of it is of a 'bean counting' nature and adds little of social value." (Stephan, 2010). These criticisms are based on the high costs that scientists sustain in applying for competitive grants considered as wasted efforts if the competition turns out to be unsuccessful.

Nonetheless, highly competitive grants require an extensive commitment in the submission phase. Scientists are asked to spend time elaborating an appealing research idea and accurately planning its execution to persuade the evaluators that they will fulfill the promised deliverables. Also, since the grant call is often designed with the requirement of having co-applicants, the application process could be an occasion for scientists to build collaboration linkages (DeFazio et al. 2009).

In this paper, we compare two groups of scientists with the same characteristics differing only in the decision to participate or not in a grant competition. Adopting a difference-in-differences approach, we assess if scientists, who decided to apply, perform differently from the

ones who did not. We use a novel dataset of 775 grant applicants to SINERGIA, a Swiss funding program sponsoring interdisciplinary collaboration where researchers are asked to submit a joint proposal to access funds. We then select a control sample of potential applicants, i.e. scientists with observable characteristics as close as possible to the applicants in our sample using a propensity score matching approach. Since the scientist's observable characteristics used to match applicants with potential applicants might not be perfect proxies for the scientist's quality, commitment, and ability, an instrumental variable approach is added to assure the reliability of our identification strategy.

We find that when applying for a SINERGIA grant, regardless of the result of the application, scientists increase their productivity in terms of number of publications and increase the average impact factor of the journals where they publish. These results suggest that the efforts incurred to apply for the grant pay in the subsequent quality and quantity of the researchers' scientific production. Applicants also expand their collaboration network by co-authoring with their co-applicants. However, in writing multi-disciplinary and long-term projects for a grant like SINERGIA, scientists enter new fields in which they have to acquire new knowledge (Azoulay et al., 2011) and where their reputation requires time to be established. As a result, we observe, for the applicants, a reduction in the average number of citations received per paper.

If on one side scientists question about the utility of participating in a grant competition, on the other side, there is rising attention of researchers for managing their budgets efficiently, partly driven by the growing desire of governments to control public money spending. As an example, illustrating the increasing public pressure on scientists, since the early nineties, the U.S. government is asking funding agencies to report the outcomes of projects publicly supported. The U.S. Government Performance and Results Act of 1993 states that "the Director of the Office of

Management and Budget shall require each agency to prepare an annual performance plan covering each program activity set forth in the budget of such agency. Such plan shall "[…] establish performance goals to define the level of performance to be achieved by a program activity; […] establish performance indicators to be used in measuring or assessing the relevant outputs, service levels, and outcomes of each program activity" (sec. 2803). As an example of an action taken to respond to such government regulatory interventions, one of the largest American evaluation programs assessing the impact of public investment in research, STAR METRICS, was launched to provide taxpayers with precise information on the value of their investments (Lane, 2011).

Despite a growing demand for an evaluation of publicly supported scientific research, extant studies do not provide convergent findings on the effect of receiving funds on researchers' scientific outcomes. The results suggest a limited impact of funding on the main scientific outcomes of scientists, but the magnitude of the effects and the outcomes analyzed vary across studies (Arora and Gambardella, 2005; Jacob and Lefgren, 2011; Gush et al., 2015; Azoulay et al., 2015; Carayol and Lanoe, 2017).

The disparity in the empirical findings could result from several technical limitations. In a comprehensive review, Jaffe (2002) identifies three main difficulties encountered when attempting to evaluate the effects of research funding. First, there might be information availability issues since it is often difficult to retrieve detailed information about the full sample of scientists applying for a grant, awarded and non-awarded, and demographic information about the studied scientist is sparse. Second, even with the accessibility to such data, the estimation of the funding effect might be biased because most productive scientists are also the ones having a higher probability to be

funded. Third, the standard bibliometric measures such as the number of the publications and the average impact factor might provide only a partial picture of the effects of being awarded a grant.

Our study proposes a set of solutions to tackle these obstacles. We use a comprehensive dataset of scientists including both awarded and non-awarded applicants and, following Fox (1983), we exploit the richness of our dataset to include both individual-level variables and environmental characteristics in our analysis. We then apply a precise matching process, and several robustness checks to account for the selection bias. We also introduce novel scientific outcomes to capture more extensive aspects of scientific production such as collaboration and learning.

In our analysis of the impact of funding, we find that receiving funds represents a proper incentive to realize the potential collaborations claimed in the application phase. Precisely, we find that the probability of co-authoring with at least one other scientist listed in the application is higher within awarded applications. However, concerning the productivity of funded researchers, we observe similar results to the ones of the literature. Being awarded has a limited but not significant impact on the quality and quantity of a researcher's scientific productivity.

The organization of this paper is as follows: Section 2 sets the empirical context, Section 3 describes the data and main variables, Section 4 exposes the estimation strategy, Section 5 presents the findings, Section 6 discusses the results, and Section 7 concludes.

2. Empirical context

The Swiss National Science Foundation (SNSF) is the main national funding agency in Switzerland. It plays in the country the same role of the National Science Foundation (NSF) in the United States or the European Research Council (ERC) in Europe. The SNSF supports researchers' activities and their careers. SINERGIA program is one of the flagship programs in its portfolio. It was launched in 2008 and designed to promote breakthrough research and collaboration of scientists affiliated with different institutions. As mentioned in the application guidelines, scientists are required to collaborate with colleagues from another institution as a condition for securing research funding, i.e., scientists need to submit a proposal for a "research work carried out collaboratively" (SNSF, 2011). The application process for SINERGIA is similar to the one of NSF and ERC grants. Researchers based in research universities and public research institutions obtain public funds on a competitive basis by submitting their proposal to the selection committee of the SNSF. The committee then selects the most promising projects to which the funds are allocated.

In most cases, a SINERGIA project involves four or five scientists led by a main proponent coordinating the overall project. All disciplines are eligible for funding through the program. Applicants propose interdisciplinary projects or projects where co-applicants belong to the same field, but are specialized in different sub-fields[1]. The criteria considered in evaluating the application are the value added by the joint research approach, the research complementarities of the applying groups, and the coherence of the projected collaboration. The screening of applications is a two-step evaluation process. In the first step, external reviewers assign a

[1] An example of a project in different disciplines is one including Math, Hydrology and Geophysics, while an example of a single-discipline project with two sub-disciplines is one with Biochemistry and Genetics.

provisional score to each application. In the second step, an internal committee of the SNSF, the Specialized Committee for Interdisciplinary Research, based in Bern, assigns a final score to each application using a scale where 6 is the highest score, and 1 the lowest. The evaluation process takes six months to be completed. Applications are funded until the annual budget quota is reached. The decision to award projects is also based on the funds available; all awarded projects received funds.

3. Data

This section describes the characteristics of applications and applicants (paragraph 3.1), illustrates the procedure applied to select a control sample of potential applicants (paragraph 3.2), and presents the scientific outcomes (paragraph 3.3).

3.1 Applications and applicants

The scarcity of information disclosed by the funding agencies about their application selection process has often limited the capacity of scholars to estimate the effects of public funding activities adopting ideal identification strategies. Our scientific partnership with SNSF provided us with the opportunity to have all –both awarded and non-awarded- grant applications submitted by Swiss researchers applying for the SINERGIA grant in the period 2008-2012[2]. We also have access to the scores assigned to the applications, the final funding decisions, and demographic information about applicants. We complement this information with applicants' publication records using the Elsevier's Scopus database. To perform our analysis, we select applications in

[2] All concerned applicants were contacted by the SNSF and had the possibility to oppose the transmission of their data.

Engineering and Science & Medicine[3]. Our final sample includes 255 grant applications and 775 distinct applicants. Our unit of analysis is the pair applicant-application. Considering that each applicant can be involved in more than one application, our sample counts 1,060 applicant-application pairs.

As application characteristics, we consider five sets of variables measuring the application funding decision, the project size, its quality, the applicant team composition, and discipline. We capture the funding decision using a dummy that equals one if the application is awarded, zero otherwise (*Awarded*). The size of the research project is proxied by two variables, the *Amount requested* in Swiss Francs (CHF) and the number of co-applicants listed in the application document (*N. of co-applicants*). The quality of the project is evaluated using a variable that ranges from 1 to 6 according to the grade assigned by the selection committee to the application (*Grade*). We proxy the ethnic composition of the applicants' team using a dummy that equals one if all the applicants are affiliated with Swiss institutions (*Swiss team*) and their geographical dispersion using a continuous variable measuring the average distance in terms of travel time between the researcher's affiliation and the co-applicants' affiliations (*Distance hours*). For gender, we use a dummy that equals one if there is at least one female researcher among the co-applicants (*At least one female researcher on the team*). Finally, we identify the discipline of the application with a dummy that equals one if the application is in the domain of Science and Medicine and zero if the application is in the domain of Engineering (*Science & Medicine*).

Table 1 reports the key figures describing application characteristics. The applications in our sample are composed, on average, by 4.19 members, with a minimum of 2 and maximum of

[3] In this study, we exclude from the original sample applications in the Humanities and Social Sciences since book contributions represent a significant part of the field publication outcomes and are not collected with accuracy in the Elsevier's Scopus database. Applications in the Humanities and Social Sciences represent 19% of the total sample.

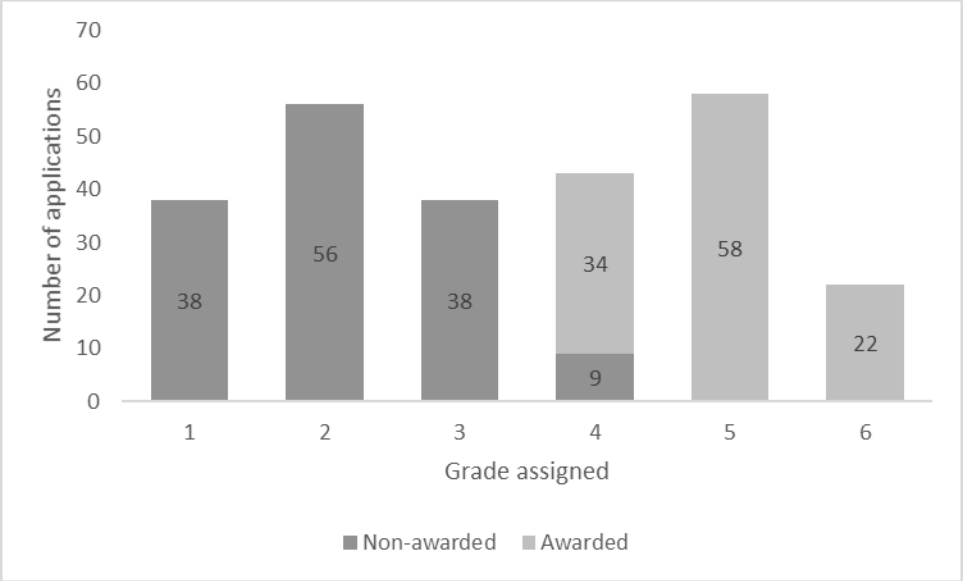
11 members. Concerning team composition, about 13% of the teams have only Swiss members, while the others are multi-national teams. When classified by discipline, 36% of the applications are in Engineering, whereas 64% are in Science & Medicine. A SINERGIA grant covers personnel costs, research costs, coordination costs and, to a limited extent, investment costs. The average amount requested per application is 1.67 million of CHF, with a minimum of 0.35 million CHF and a maximum of 6.85 million CHF. Figure 1 represents the distribution of the number of grant applications by the score assigned and the final funding decision. A total of 9% of the applications obtained the maximum score, 6, and 45% of the applications were awarded.

Table 1: Application characteristics (Number of applications=255).

	Mean	Std. Dev.	Min	Max
Awarded	0.45	0.50	0	1
Amount requested (in million CHF)	1.67	0.76	0.35	6.85
N. of co-applicants	4.19	1.59	2	11
Grade	3.39	1.60	1	6
Swiss team	0.13	0.33	0	1
At least one female researcher on the team	0.36	0.41	0	1
Distance hours	4.39	3.59	1	19.75
Science & Medicine	0.64	0.48	0	1

The table shows the key figures concerning the 255 grant applications included in our sample. The table reports mean, standard deviation, minimum and maximum values of five sets of variables regarding the application funding decision (*Awarded*), the project size (*Amount requested* and *N. of co-applicants*), its quality (*Grade*), the applicant team composition (*Swiss team*, *At least one female researcher on the team*, and *Distance hours*), and the main discipline of the application (*Science & Medicine*).

Figure 1: Distribution of the number of grant applications by the score assigned and final funding decision.



The figure shows the distribution of the grant applications according to the grade assigned by the evaluation committee. The grades range from 1 to 6, where 6 is the highest grade.

As applicant characteristics, reported in Table 2, we consider two sets of variables, respectively measuring the applicant's demographic and bibliometric characteristics before the application year. The demographic characteristics include a dummy that equals one if the researcher is a female and zero otherwise (*Female*) and the variable *Seniority* that measures the time since the first year of scientific activity of the researcher[4]. We consider a set of bibliometric measures proxying the applicant's publication characteristics before the application year. These measures include a set of variables computed in the five years preceding the application: the publication count (*Publication count pre-application*), the average impact factor of the journals where the applicant published (*Average IF pre-application*), the average number of citations received per paper (*Average citations pre-application*), the yearly average number of authors per

[4] To identify this beginning year, we track the first publication authored by the researcher looking at her self-citations within the publications available in our study sample (2003-2015). We consider this publication year as a proxy for the year when the researcher started her activity.

paper (*Average authors pre-application*), and the existence of at least one co-authored paper between the scientists and the other applicants (*Co-applicant collaboration pre-application*).

The SINERGIA funding program targets established researchers who demonstrated their ability to conduct excellent quality independent research. In most cases, applicants are associate or full professors with good publication records. The average seniority of the applicants is 18.52 years since the start of their research activity. The average number of applicants' publications is 31.35 in the five years preceding the application year. An average applicant has received 4.28 citations per paper at the application time and has published on journals with an average impact factor of 5.59. When looking at gender distribution, 15% of applicants in our study sample are female. In 40% of the cases, the applicant researchers have established collaborations with the other applicants, i.e., they have co-authored at least one article with them in the five years preceding the application.

Table 2: Applicant characteristics (Number of applicants=775).

	Mean	Std. Dev.	Min	Max
Seniority	18.52	9.37	0.00	53.00
Female	0.15	0.36	0.00	1.00
Publication count pre-application	31.35	27.29	1.00	225.00
Average IF pre-application	5.59	3.60	0.10	28.61
Average citations pre-application	4.28	4.26	0.04	48.62
Average authors pre-application	5.14	1.13	1.00	10.40
Co-applicant collaboration pre-application	0.40	0.47	0.00	1.00

The table shows the key figures concerning the 775 grant applicants included in our sample. It reports the mean, standard deviation, minimum and maximum values of two sets of variables regarding the applicants' demographic characteristics (*Seniority* and *Female*) and the applicants' bibliometric characteristics (*Publication count pre-application, Average IF pre-application, Average citations pre-application, Average authors pre-application,* and *Co-applicant collaboration pre-application*).

3.2 Searching for a group of potential applicants

To estimate the effect of applying for a grant we construct a control sample of researchers who would have been eligible to apply but did not apply for a SINERGIA grant. To do so, we retrieve a group of potential applicants with profiles similar to the ones of the applicants of our sample. We find a potential applicant, i.e., a matched control, for each of the 1,060 applicant-application pairs in our sample. To identify potential applicants, we proceed in two steps. First, we define a large pool of scientists eligible to apply for SINERGIA. Second, we extract from this pool of scientists the ones who match the profiles of the applicants using a propensity score matching approach.

We consider as eligible scientists all the publishing scientists affiliated with one of the twelve major Swiss universities[5]. From all the publications of the scientists affiliated with those universities in the period 2003-2015, we retrieve 25,715 authors who were active in the period 2008-2012, i.e., the period during which the SINERGIA grants were awarded. We consider a scientist active in a given year *t* if she has at least one publication in the time window *[t-5,t-1]* and at least one in *[t,t+4]*. Each of the 25,715 scientists is observed yearly leading us to a pool of 86,694 scientist-year pairs.

To extract from the pool of scientist-year pairs the most appropriate control for each applicant-application pair we use a propensity score matching based on a logit estimation of the probability of applying for SINERGIA[6]. In this estimation, the dependent variable (*Applicant*)

[5] University of Neuchatel, ETHZ, EPFL, University of Lausanne, University of Fribourg, University of Genève, University of Bern, University of Basel, University of Lugano, University of Zurich, University of Luzern, and University of St. Gallen.

[6] To be conservative, and exclude the possibility that some of our results are driven by some loosely matched controls, we also limit the analysis to the controls having a propensity score different by less than 1% from the one of the actual applicant they are matched with. Our results remain stable across the two approaches.

equals one for the 1,060 applicant-application pairs and zero for all the remaining scientist-year pairs in the pool. We identify 1,060 controls, one for each of the 1,060 applicant-application pairs. We define the 1,060 controls as the potential applicants to the SIENRGIA grant.

To identify potential applicants, we consider as relevant matching characteristics: the researcher's *Seniority*, her fundraising profile and her bibliometric characteristics before the application year[7]. The fundraising profile of a scientist is captured through two variables: *Other active funding* and *Previous expired funding*. The variable *Other active funding* is a dummy which equals one if the scientist has at least one active project granted (other than SINERGIA) at the moment of the application and zero otherwise, while the variable *Previous expired funding* is a dummy which equals one if the scientist has raised funds in the past with a grant that was expired at the moment of the application to SINERGIA and zero otherwise. As funding, we consider the European Union grants[8] and the SNSF grants other than SINERGIA[9].

The bibliometric characteristics include the variables *Publication count pre-application*, *Average citations pre-application, Average IF pre-application, Average authors pre-application*. To improve the matching quality, we include in the regression the average yearly variation of the bibliometric characteristics. These additional variables allow us to take also into account the trends of these bibliometric indicators over the 5-year window considered. Precisely, we calculate the average yearly growth (decline) of the publication count over the five years of observation preceding the application year (*Average publication trend*). Similarly, we construct the variable

[7] For a potential applicant, we consider as application year the year when her matched applicant submitted her application.

[8] From the European Union we retrieved the data of the grants awarded between 1998 and 2013, namely FP5 (1998-2002), FP6 (2002-2006) and FP7 (2007-2013). We collected this data on the CORDIS platform online (https://data.europa.eu).

[9] For Switzerland, we used the P3 database of the Swiss National Science Foundation (SNSF) which makes data available on the projects and people that it has supported (http://p3.snf.ch).

Average citation trend, *Average IF trend*, and *Average authors trend*. To account for the continuity in the productivity of the scientists, we add as further bibliometric characteristic the variable *Productivity break* which counts the number of years without any publication in the five-year window preceding the application year (Mairesse and Pezzoni, 2015).

We also include as relevant matching characteristic the stock of knowledge of the scientist before the application year. The stock of knowledge, as more extensively described in paragraph 3.3, is represented by the number of distinct journals listed in the references of the articles published by the scientist in the five years preceding the application (*N. of journals pre-application*).

Finally, we control for *Application year*, *Affiliation*, and *Discipline*[10] fixed effects. Table 3 shows the logit estimation of the regression used to predict the matching probability.

[10] See Section A in the appendix for a detailed description of the attribution of disciplines to scientists.

Table 3: Propensity score matching regression, logit estimation.

	Logit Applicant
Seniority	0.00015
	(0.0049)
Other active funding	1.13***
	(0.10)
Previous expired funding	0.16
	(0.10)
Publication count pre-application	0.018***
	(0.0023)
Average IF pre-application	0.079***
	(0.015)
Average citations pre-application	-0.13***
	(0.017)
Average authors pre-application	0.71***
	(0.023)
Average publication trend	0.040
	(0.033)
Average IF trend	-0.044
	(0.036)
Average citation trend	-0.023
	(0.033)
Average authors trend	-0.045
	(0.057)
Productivity break	-1.49***
	(0.073)
Log(N. of journals pre-application)	0.74***
	(0.084)
Constant	-8.27***
	(0.40)
Dummy Application year	Yes
Dummy Discipline	Yes
Dummy Affiliation	Yes
N. of Scientists	25,715
Observations	86,694
Pseudo R-squared	0.60

The table shows the coefficients estimated for the regression predicting the probability of applying for a SINERGIA grant as a function of a scientist's demographic and bibliometric characteristics. The regressors include fixed effects for *Application year*, *Affiliation*, and *Discipline*. The dependent variable is a dummy that equals one if the scientist applies to SINERGIA, zero otherwise. The 86,694 observations refer to the entire pool of scientist-year pairs from which we extract 1,060 potential applicants, one for each 1,060 applicant-application pairs. We estimate a Logit model. In reporting the statistical significance of the coefficients, we apply the standard thresholds i.e. *** $p<0.01$, ** $p<0.05$, * * $p<0.1$.

Table 4 shows the descriptive statistics for four groups of researchers: awarded applicants, non-awarded applicants, the entire set of applicants, and the potential applicants selected with the propensity score matching approach. Statistical t-tests on the averages of the main bibliometric characteristics of applicants (column 3) and potential applicants (column 4) do not reject the null

hypothesis that the two groups have the same averages. This evidence confirms the accuracy in the selection of potential applicants with profiles similar to the applicants. The only case where we reject the null hypothesis of the t-test is for the variable *N. of journals pre-application*[11].

Table 4: Descriptive statistics for the four groups of scientists before the application year.

	(1) Awarded (469 obs.)			(2) Non-Awarded (591 obs.)		
	Mean	Min	Max	Mean	Min	Max
Publication count pre-application	30.23	1.00	225.00	34.34	2.00	183.00
Average IF pre-application	6.14	0.10	19.97	5.19	0.10	28.61
Average citations pre-application	4.61	0.12	38.67	4.03	0.04	48.62
Co-applicant collaboration pre-application	0.37	0.00	1.00	0.42	0.00	1.00
N. of journals pre-application	125.35	1.00	495.00	135.50	1.00	594.00
Average authors pre-application	5.26	1.00	10.40	5.08	1.33	10.16
Seniority	17.37	0.00	50.00	17.95	0.00	52.00
Other active funding	0.48	0.00	1.00	0.38	0.00	1.00
Previous expired funding	0.37	0.00	1.00	0.35	0.00	1.00
Productivity break	0.50	0.00	4.00	0.41	0.00	4.00
	(3) Applicants (1,060 obs.)			(4) Potential Applicants (1,060 obs.)		
	Mean	Min	Max	Mean	Min	Max
Publication count pre-application	32.52	1.00	225.00	30.59	1.00	191.00
Average IF pre-application	5.61	0.10	28.61	5.51	0.10	35.21
Average citations pre-application	4.29	0.04	48.62	4.18	0.01	32.24
Co-applicant collaboration pre-application	0.40	0.00	1.00	0.02	0.00	1.00
N. of journals pre-application	131.01	1.00	594.00	114.50	2.00	515.00
Average authors pre-application	5.16	1.00	10.40	4.87	1.25	15.00
Seniority	17.69	0.00	52.00	17.93	0.00	52.00
Other active funding	0.43	0.00	1.00	0.46	0.00	1.00
Previous expired funding	0.36	0.00	1.00	0.40	0.00	1.00
Productivity break	0.45	0.00	4.00	0.54	0.00	4.00

The table shows the average bibliometric characteristics before the application year for four groups of scientists: Awarded applicants (469 obs.), non-awarded applicants (591 obs.), whole sample of the applicants (1,060 obs.), and potential applicants selected according to the propensity score matching procedure (1,060 obs.). Comparing the characteristics of Applicants and Potential applicants (Columns 3 and 4), we find no statistical difference, i.e. p-value>0.1, of the means of the two groups in terms of *Publication count pre-application, Average IF pre-application, Average citations pre-application, Average authors pre-application, Seniority, Other active funding, Previous expired funding,* and *Productivity break.*

3.3 Scientific outcomes

[11] To test the robustness of our analysis, we constructed an alternative control sample for which *N. of journals pre-application* is the main matching criterion. Our results remain stable across the two control samples. Results of this robustness test are available upon request.

We measure scientists' outcomes along five dimensions. The first three dimensions are standard bibliometric measures commonly used in the literature. We then add two novel measures to draw a more extensive picture of the outcomes scientists can achieve in their scientific research. All these variables are computed in the five-year window after the SINERGIA application year.

In assessing publication quantity and quality of scientists, we use standard bibliometric measures: publication count, citations and impact factor. Following the mainstream literature (Arora and Gambardella 2005; Carayol and Lanoe 2017), we account for the publication quantity as the count of papers published by a scientist (*Publication count*). As for publication quality, we use impact factor (*Average IF*) and citations (*Average citations*) as two quality measures capturing different aspects of the impact of a publication. Citations indicate the number of times other scientists refer to a work and could be defined as the "peer assessments of the significance of scientists' contributions" (Cole and Zuckerman, 1984, pp. 231). Impact factor expresses the quality of the journal where a work is published (Long, 1992). For a work, being published in a journal with a high impact factor is commonly seen as a signal of high intrinsic quality.

When measuring the impact of a scientific grant with the standard bibliometric outcomes, as stated by Jaffe (2002), we are only "examining this tiny piece of a very complicated puzzle […], we are not looking at the spillovers that are perhaps the true reason for these programs." In our case, collaboration and knowledge transfer are stated as "true reasons" for implementing the grant. We consider two novel measures aiming to capture these two core aspects of the SINERGIA grant: successful collaboration and individual learning.

To measure successful scientific collaborations, we consider the work relationship consolidated among SINERGIA co-applicants through a co-authorship. Specifically, we measure

a successful collaboration as a dummy (*Co-applicant collaboration*) that equals one if a scientist co-authors at least one paper with her co-applicants and equals zero if she does not.

To measure the knowledge transfer among scientists, we construct a novel measure of individual learning. As stated by Huber (1991), "Learning consists in knowledge acquired by any unit of an organization and available for acting upon." Hence, building on Ayoubi et al. (2017), we define individual learning as the increment to the stock of knowledge of a scientist between two periods. Following Uzzi et al. (2013), we use the journals cited as the building blocks of knowledge and thus consider the knowledge stock of a scientist at a given point in time as the journals that she cited in her publications. Our proxy for individual learning is then the difference between the knowledge stock of the scientist before and after the application time. Specifically, we measure individual learning at a given moment as the number of distinct journals in the references of the publications of the scientist not observed in a previous period (*Learning*).

Table 5 shows the descriptive statistics of the scientist's outcomes for the applicants (distinguishing awarded and not awarded) and the potential applicants selected.

Table 5: Descriptive statistics for the scientist's outcomes after the application year.

	(1) Awarded (469 obs.)			(2) Non-Awarded (591 obs.)		
	Mean	Min	Max	Mean	Min	Max
Publication count	36.28	1.00	189.00	38.64	2.00	251.00
Average IF	5.95	0.73	20.48	4.97	0.10	19.78
Average citations	2.82	0.03	17.38	2.36	0.06	27.55
Co-applicant collaboration	0.71	0.00	1.00	0.61	0.00	1.00
Learning	95.46	1.00	368.00	101.74	2.00	556.00
	(3) Applicants (1,060 obs.)			(4) Potential Applicants (1,060 obs.)		
	Mean	Min	Max	Mean	Min	Max
Publication count	37.60	1.00	251.00	30.39	1.00	165.00
Average IF	5.40	0.10	20.48	5.21	0.10	31.96
Average citations	2.56	0.03	27.55	3.93	0.11	78.13
Co-applicant collaboration	0.66	0.00	1.00	0.01	0.00	1.00
Learning	98.96	1.00	556.00	70.22	1.00	318.00

The table shows the descriptive statistics of the scientific outcomes after the application year for four groups of scientists: Awarded applicants (469 obs.), non-awarded applicants (591 obs.), whole sample of the applicants (1,060 obs.), and potential applicants selected according to the Propensity Score Matching procedure (1,060 obs.). The five variables included in the table, *Publication count*, *Average IF*, *Average citations*, *Co-applicant collaboration*, and *Learning* are the dependent variables of the regression models estimated in Section 5.

4. Methodology

4.1 Estimation strategy

To estimate both the effects of applying for a grant and receiving the funds, we rely on a difference-in-differences approach where we compare the changes in scientific outcomes between before and after applying or receiving the funds (Angrist and Pischke, 2008). The identifying assumption is that a scientist's outcome trends would be the same in the absence of the treatment represented by applying for the grant program and receiving the funds[12], repectively.

4.2 Effect of applying

To formally evaluate whether applying has a significant effect on the subsequent scientist's outcomes we compare the applying scientists to the matched control sample of potential applicants described in paragraph 3.2 above.

For estimating the application effect, we rely on the equivalent formulation of the difference-in-differences reported in Equation 1 and estimated with an Ordinary Least Squares (OLS).

$$Scientist's\ outcome_{it}=\beta_0+\beta_1\ Applicant_i + \beta_2\ Post\text{-}Application_{it} + \beta_3(Applicant_i*Post\text{-}Application_{it}) + (Scientist's\ characteristics_i)'\ \beta_4 +\varepsilon_{it}$$

(Equation 1)

Where *i* and *t* refer to the scientist *i* observed at time *t.* We observe the scientist in two periods, before her application (*t=0*) and after her application (*t=1*). To define the two periods for a

[12] We formally test for this assumption using the fully flexible model for parallel paths introduced by Mora and Reggio (2017) and find that our sample fulfills this identifying assumption. Results are available upon request.

potential applicant who, by definition, does not apply to SINERGIA, we use the application year of her matched applicant.

The variable *Scientist's outcome*$_{it}$ is, in turn, one of the five dependent variables described in paragraph 3.3 (*Publication count*, *Average IF*, *Average citations*, *Co-applicant collaboration*, *Learning*). All the dependent variables, with the exception of *Co-applicant collaboration*, are log-transformed[13].

The dummy *Applicant*$_i$ equals one if scientist *i* is an applicant to SINERGIA and zero otherwise. *Post-Application*$_{it}$ is a time dummy that takes a value of zero if we observe the scientist *i*'s outcomes before her application (*t=0*), and a value of one if we observe the scientist's outcomes after (*t=1*). The interaction term *Applicant*$_i$**Post-Application*$_{it}$ marks a scientist *i* who experienced the application and whose outcomes are observed after the application time. The estimated coefficients of this interaction measure the effect of applying for a SINERGIA grant[14].

The vector *Scientist's characteristics* includes the variable *Seniority* as well as the *Application year*, *Affiliation*, and *Discipline* fixed effects.

A potential issue with our results could come from the endogeneity arising due to applicants' self-selection. The propensity score matching along with the difference-in-differences approach described in this section are meant to take into account this concern. In Section B of the appendix, we present an additional estimation exercise using an instrumental variable approach

[13] Note that this log transformation is possible since all the values of the variables of interest are strictly positive for all the scientists of our sample (See minimum values in Table 5).

[14] In comparing applicant versus potential applicant scientists, the heterogeneity of the former category might bias our results. Specifically, applicant scientists include awarded and non-awarded scientists that might behave in a different way. To address this concern, we run a robustness check where we include as controls the variable *Awarded* and the interaction term *Awarded*Post-Application*. The new results, available upon request, are consistent with the results reported in Table 6.

evaluating the impact of applying for a grant (Wooldridge, 2012). Our main results remain stable across the different econometric exercises.

4.3 Effect of being awarded

After having evaluated the impact of applying, we focus our attention on the subsample of applicant scientists, and we estimate the effect of being awarded a SINERGIA following Equation 2.

$$Scientist's\ outcome_{it}=\beta_0+\beta_1\ Awarded_i+\beta_2\ Post\text{-}Application_{it}+\beta_3(Awarded_i*Post\text{-}Application_{it})+(Scientist's\ characteristics_i)'\ \beta_4+(Application\ characteristics_i)'\ \beta_5+\varepsilon_{it}$$

(Equation 2)

Where *i* and *t* refer to the scientist *i* observed at time *t.* We observe the scientist in two periods, before the application (*t=0*) and after the application (*t=1*). The variable *Scientist's outcome*$_{it}$ is, in turn, one of the five dependent variables described in paragraph 3.3 above (*Publication count*, *Average IF*, *Average citations*, *Co-applicant collaboration*, *Learning*). All the dependent variables, with the exception of *Co-applicant collaboration*, are log-transformed[15].

The dummy *Awarded*$_i$ equals one if scientist *i* was awarded a SINERGIA grant and zero otherwise. *Post-Application*$_{it}$ is a time dummy that takes a value of 0 if we observe the scientist *i*'s outcomes before the application (*t=0*), and a value of 1 if we observe the scientist's outcomes after

[15] Although some previous studies apply a Regression Discontinuity Design (RDD) approach (Jacob and Lefgren, 2011), this approach is not suitable in our case. The main reason is that we do not have a ranking of the applications but six grades, from 1 to 6. For us, the threshold is located at grade 4 since all the applications above it are awarded while none of the ones below are. We cannot rank the applications within grade 4 to position correctly the threshold. This limitation is related to the procedure used by SNSF to select the awarded applications graded 4. All the application graded 4 go through a second round of evaluation where the committee revise each application but without producing a ranking. In order to test the robustness of our results with an approach similar to the RDD, we conduct our analysis by restricting our sample to the applications graded 4. The results, available upon request, are consistent with the results reported in Table 7.

($t=1$). The interaction term $Awarded_i*Post\text{-}Application_{it}$ marks a scientist i who was awarded a grant and whose outcomes are observed after the application time. The estimated coefficients of this interaction measure the effect of being awarded a SINERGIA grant.

The vector *Scientist's characteristics* includes the variables *Seniority*, *Female*, *Other active funding*, and *Previous active funding* as well as the *Application year*, *Affiliation*, and *Discipline* fixed effects. The vector *Application characteristics* includes the variable *Grade*, the dummies *Swiss Team*, *At least one female researcher*, *Science & Medicine*, and the continuous variables *Log(Amount requested)*, *Log(N. of co-applicants)*, *Log(N. of disciplines)*, and *Log(1+Distance hours)*. We also control for the presence of previous applications to SINERGIA and previous awarded SINERGIA applications using two dummies, *Previous SINERGIA application* and *Previous SINERGIA awarded*. The first variable equals one if the applicant has at least one previous application to SINERGIA and the second equals one if the applicant has at least one previous SINERGIA awarded, zero otherwise.

5. Results

This section summarizes the results of the two main regressions described in Section 4. First, we present the main findings for the difference-in-differences regression estimating the effect of applying for the SINERGIA grant (Equation 1)[16]. Second, we present the results of the regression estimating the impact of receiving the funds (Equation 2).

Table 6 reports the impact of applying for a SINERGIA grant. Columns 1 to 3 cover the regression results for the three standard bibliometric measures: Publication count in logarithmic terms (*Log(Publication count)*) in column 1, the average impact factor of the journals where the scientist publishes in logarithmic terms (*Log(Average IF)*) in column 2, and the average number of citations received per paper in logarithmic terms (*Log(Average citations)*) in column 3. Column 4 and 5 report the results for the dummy co-applicants' collaboration realized (*Co-applicant collaboration*) and the individual learning of scientist in logarithmic terms (*Log (Learning)*), respectively.

We find that scientists who applied for a SINERGIA grant are more productive in quantitative terms than scientists who did not apply. Specifically, applicants publish on average 43% papers more than non-applicant in the five-year window following the application. We also observe that they increase the average impact factor of the journals where they publish by 7%. However, applicants' articles receive, on average, 33% fewer yearly citations than those of the potential applicants. Applicants have a 19% higher probability of establishing a co-authorship with their co-applicants than potential applicants and learn more on average.

[16] For the two discrete variables *Publication count* and *Learning* we also performed a Poisson estimation. The results of these estimations, available upon request, are consistent with the results reported in Table 6.

Table 6. Regression results for the estimation of Equation 1 comparing applicants to potential applicants.

	(1) OLS	(2) OLS	(3) OLS	(4) Probit	(5) OLS
	Log(Publication count)	Log(Average IF)	Log(Average citations)	Co-applicant Collaboration	Log(Learning)
Applicant*Post-Application	0.43***	0.070***	-0.33***	0.19***	0.36***
	(0.032)	(0.026)	(0.031)	(0.046)	(0.038)
Applicant	0.23***	-0.030	-0.11***	0.43***	0.043
	(0.032)	(0.033)	(0.041)	(0.027)	(0.032)
Post-Application	-0.25***	-0.067***	-0.15***	-0.013	-0.62***
	(0.027)	(0.024)	(0.029)	(0.030)	(0.028)
Seniority	0.031***	-0.0019	-0.00053	-0.0014	0.012***
	(0.0019)	(0.0014)	(0.0016)	(0.00088)	(0.0014)
Other active funding	0.33***	0.19***	0.17***	-0.0081	0.25***
	(0.033)	(0.025)	(0.031)	(0.016)	(0.028)
Previous expired funding	0.22***	-0.026	-0.068**	0.030**	0.098***
	(0.037)	(0.026)	(0.034)	(0.015)	(0.029)
Constant	1.76***	1.17***	0.99***		3.37***
	(0.080)	(0.067)	(0.075)		(0.075)
Dummy Application year	Yes	Yes	Yes	Yes	Yes
Dummy Discipline	Yes	Yes	Yes	Yes	Yes
Dummy Affiliation	Yes	Yes	Yes	Yes	Yes
Appl./Potential appl.	2,120	2,120	2,120	2,120	2,120
Observations	4,240	4,240	4,240	4,240	4,240
R-squared / Pseudo R-squared	0.435	0.261	0.243	0.39	0.466

The table shows a difference-in-differences estimation in the equivalent regression formulation, where the effect of the treatment, i.e. applying for a SINERGIA grant, can be read in the coefficient of the interaction variable *Applicant*Post-Application.* The five scientific outcomes considered are *Publication count, Average IF, Average citations, Co-applicant collaboration,* and *Learning.* The sample includes 4,240 observations. The 1,060 Applicant-application pairs and the 1,060 Potential applicants, i.e. 2,120 Applicant-application pairs/Potential applicants-matched application pairs, are observed in two periods, namely before and after the application year. Columns 1, 2, 3, and 5 report OLS estimates, whereas Column 4 reports the marginal effect of a Probit estimation that considers the binary nature of the dependent variable *Co-applicant collaboration.* Robust standard errors are clustered around the application. In reporting the statistical significance of the coefficients, we apply the standard thresholds, namely *** p<0.01, ** p<0.05, * * p<0.1.

Table 7 shows the results of the estimation of Equation 2 comparing awarded applicants to non-awarded applicants. On average, awarded applicants do not perform better than non-awarded ones regarding the quantity and quality of their scientific production and learning, but they have a greater chance to establish a co-authorship with their co-applicants. Specifically, being awarded the funds increases the probability of co-authoring with co-applicants by 17% on average.

Table 7. Regression results for the estimation of Equation 2 comparing awarded to non-awarded scientists.

	(1) OLS Log(Publication count)	(2) OLS Log(Average IF)	(3) OLS Log(Average citations)	(4) Probit Co-applicant collaboration	(5) OLS Log(Learning)
Awarded*Post-Application	0.034	0.027	0.086	0.17***	-0.025
	(0.043)	(0.026)	(0.054)	(0.042)	(0.064)
Awarded	-0.15**	-0.015	-0.11	-0.15**	-0.19**
	(0.076)	(0.083)	(0.10)	(0.069)	(0.086)
Post-Application	0.16***	-0.0087	-0.53***	0.20***	-0.25***
	(0.029)	(0.017)	(0.036)	(0.028)	(0.039)
Scientist's characteristics					
Seniority	0.023***	0.0039**	0.0058**	-0.0032*	0.012***
	(0.0024)	(0.0018)	(0.0025)	(0.0017)	(0.0021)
Female	-0.11**	0.032	0.10*	0.0080	0.033
	(0.052)	(0.045)	(0.054)	(0.040)	(0.053)
Other active funding	0.15***	0.13***	0.16***	-0.024	0.10***
	(0.039)	(0.033)	(0.043)	(0.030)	(0.037)
Previous expired funding	0.15***	-0.054	-0.083*	0.053*	0.048
	(0.043)	(0.033)	(0.045)	(0.031)	(0.037)
Application characteristics					
Grade	0.040	0.048*	0.068**	0.042**	0.071**
	(0.024)	(0.026)	(0.031)	(0.021)	(0.029)
Swiss team	-0.048	-0.0071	-0.039	-0.0028	0.095
	(0.074)	(0.063)	(0.061)	(0.066)	(0.059)
At least one female researcher	-0.047	-0.013	0.016	0.014	-0.011
	(0.042)	(0.043)	(0.048)	(0.035)	(0.049)
Log(Amount Requested)	0.028	0.12**	0.11	-0.097*	-0.013
	(0.063)	(0.060)	(0.081)	(0.050)	(0.073)
Log(N. of co-applicants)	0.051	-0.037	0.015	0.30***	0.035
	(0.067)	(0.075)	(0.094)	(0.067)	(0.098)
Log(N. of disciplines)	-0.016	0.049	-0.0081	-0.021	0.084*
	(0.037)	(0.040)	(0.048)	(0.027)	(0.051)
Science & Medicine	-0.13**	0.41***	0.42***	-0.095	0.57***
	(0.066)	(0.070)	(0.084)	(0.061)	(0.079)
Log(1+Distance hours)	0.039	0.082**	0.063*	-0.023	0.054*
	(0.027)	(0.032)	(0.037)	(0.026)	(0.030)
Previous SINERGIA application	0.13**	-0.064	-0.066	0.060	0.14**
	(0.053)	(0.051)	(0.069)	(0.046)	(0.054)
Previous SINERGIA awarded	0.067	-0.025	-0.095	0.083	0.052
	(0.076)	(0.077)	(0.089)	(0.063)	(0.066)
Constant	1.82**	-0.96	-1.22		3.18***
	(0.88)	(0.82)	(1.14)		(0.98)
Dummy Application year	Yes	Yes	Yes	Yes	Yes
Dummy Discipline	Yes	Yes	Yes	Yes	Yes
Dummy Affiliation	Yes	Yes	Yes	Yes	Yes
Applicant-application pairs	1,060	1,060	1,060	1,060	1,060
Observations	2,120	2,120	2,120	2,120	2,120
R-squared / Pseudo R-squared	0.373	0.435	0.371	0.12	0.504

The table shows a difference-in-differences estimation in the equivalent regression formulation, where the effect of the treatment, i.e. being awarded a SINERGIA grant, can be read in the coefficient of the interaction variable *Awarded*Post-Application.* The controls include fixed effects for the *Application year, Affiliation,* and *Discipline* of the scientist. The five scientific outcomes considered are the *Publication count, Average IF, Average citations, Co-applicant collaboration,* and *Learning.* The sample includes 1,060 Applicant-application pairs. The 1,060 pairs are observed before and after the treatment year, for a total of 2,120 observations. Columns 1, 2, 3, and 5 report OLS estimates, whereas Column 4 reports the marginal effect of a Probit estimation that considers the binary nature of the dependent variable *Co-applicant collaboration*. Robust standard errors are clustered around the application. In reporting the statistical significance of the coefficients, we apply the standard thresholds, namely *** p<0.01, ** p<0.05, * * p<0.1.

6. Discussion

In a period of increased selectivity of public grants, scientists are concerned about the utility of spending energy and time in participating in grant competitions where the odds to get awarded are low. In Section 5, we find that scientists who decide to apply for a grant and put the effort necessary to do so, perform differently from the ones who did not. Specifically, we find that participating in a research grant competition per se has a positive effect on the scientist's number of publications, the average impact factor of the journals where she publishes, her probability of collaborating with co-applicants and her learning. However, applicants receive on average fewer citations per paper than non-applicants.

These results suggest that the sunk costs incurred by applicants to craft their application proposal are not fruitless. The application process is time-consuming and requires dedication in writing a valuable proposal. The efforts provided to put together the project and set the research agenda boost the level of advancement and the quality of applicant's research, hence positively stimulating the subsequent number of scientific publications and the average impact factor of the journals where they are published. When applying for grants, scientists design projects and create work ties with co-applicants. We find that they build on these ties afterwards regardless of the result of the grant competition. Applicants are exposed to the knowledge of their co-applicants and spillovers are likely to occur. Hence, the observed positive impact of applying on learning could be explained by the interactions with their co-applicants when crafting the project (Ayoubi et al. 2017).

Interestingly, we observe that applicants to SINERGIA receive fewer citations on average than potential applicants. This observed decrease in citations could be the result of a decline in the quality of the scientific production or a loss of visibility due to entry to new fields of research when

scientists are involved in interdisciplinary projects (Azoulay et al. 2014). The observed increase in the average impact factor seems to discard the first interpretation, i.e. a decline in quality. The decline in citations appears hence driven by an entry to research fields with which applicants are not familiar and where they need time to establish a reputation. Entering new research fields, applicants need to acquire new knowledge. In Table 6 (Column 5) we find that applying stimulates the tendency of scientists to learn. The lack of visibility hypothesis would be then confirmed if we observe that the knowledge acquired shows a move towards new fields of research.

To give empirical ground to the hypothesis of the decreased visibility when entering new fields, we implement a test to evaluate how different is the newly acquired knowledge compared to the initial stock of knowledge of the applying scientists. Building on Ayoubi et al. (2017), we measure this difference using the cognitive distance separating the scientist's knowledge stock before the application and the knowledge acquired after the application. Specifically, we estimate the average cognitive distance between the newly acquired knowledge and the original knowledge stock of the scientist by using a measure based on the list of journals the scientist references in her scientific publications. We consider these journals as a proxy for a scientist's knowledge mobilized in her research[17].

Table 8 shows that the knowledge newly acquired by applicants is further than for potential applicants. In other terms, applying for a SINERGIA grant seems to encourage scientists to enter new fields of knowledge. This last result holds regardless of the outcome of the grant competition since being awarded has no significant effect on the entry to new fields.

[17] For a more in-depth discussion of the process and the tools used, see Section C of the appendix.

Table 8: Regression testing entry to new fields of knowledge.

	OLS Log(1+Journal distance)
Applicant	0.22***
	(0.042)
Awarded	-0.065
	(0.050)
Seniority	0.0083***
	(0.0017)
Other active funding	0.0047
	(0.044)
Previous expired funding	0.14***
	(0.036)
Constant	6.17***
	(0.11)
Dummy Application year	Yes
Dummy Discipline	Yes
Dummy Affiliation	Yes
Appl./Potential appl.	2,120
R-squared	0.224

The table shows the estimated effects of applying (*Applicant*) and being awarded a grant (*Awarded*) on the scientist's entry to new fields of knowledge. The entry to new fields of knowledge is proxied by the average cognitive distance between the newly acquired knowledge and the original knowledge stock of the scientist relying on a measure based on the list of journals cited in the scientist's articles (*Journal distance*). We observe 2,120 Applicant-application pairs/Potential applicants-matched application pairs, of which 1,060 are Applicant-application pairs, and 1,060 are Potential applicants. Robust standard errors are clustered around the application. In reporting the statistical significance of the coefficients, we apply the standard thresholds, namely *** p<0.01, ** p<0.05, * * p<0.1.

Having assessed the value of applying for a competitive grant for scientists, we consider the impact of the funding decision for applicants. Doing so, we contribute to the scarce literature on the effect of receiving funds on researchers' scientific outcomes. We find that being awarded has no significant effect on the scientific productivity of grant recipients, regarding the quantity or quality of papers published. When considering the values of the estimated coefficients in Table 7, finding little and not statistically significant impact on publication productivity, we obtain results consistent with the literature on the impact of funding on research (Arora and Gambardella, 2005; Jacob and Lefgren, 2011; Gush et al., 2015; Azoulay et al., 2015; Carayol and Lanoe, 2017). However, concerning our novel measures, we observe that co-applicants of an awarded project have greater chances to co-author with each other than co-applicants of a non-awarded project. In other words, financial resources have no impact on the individual publication productivity but provide an incentive to consolidate the collaboration started in the application phase. We might

explain these results on productivity by the fact that scientists are not strictly dependent on public grants to sponsor their research. They might have access to private funding or use their public bulk funding. However, if a grant is released, scientists are incentivized to finalize the collaborations planned when they submitted their projects.

7. Conclusion

In this paper, we find that scientists participating in a research grant competition reap the benefits of the efforts spent, even if they do not obtain the desired reward. Applicants for a SINERGIA grant increase their scientific productivity and learn while exploring new fields of knowledge. These positive externalities lead us to claim that "the important thing is not to win; it is to take part."

This noteworthy result contributes to the debate in the scientific community concerning the utility for researchers to spend time writing proposals to raise money for their research. Our results suggest that scientists derive some benefits from the time spent writing proposals. Hence, scientists should be less reluctant in investing time and effort entering research grant competition since these could be the occasion to launch new trends of research, build working ties with fellow researchers and acquire new knowledge. On the side of funding agencies, our results imply that promoting the calls for proposals as well as encouraging scientists to apply could be as efficient as increasing the funds dedicated to financing research projects.

During the exploratory phase of our work, discussing directly with the SNSF and with actual applicants for SINERGIA, we noted that a peculiarity of the SINERGIA grant is the fact that the administrative requirements of the submitted proposals are limited compared to other grants. In other words, most of the work scientists do when applying for the grant is directly related

to the scientific project they are crafting and could, therefore, be useful in further research regardless of the result of the competition. This characteristic of SINERGIA could be partly responsible for our findings and it, more importantly, suggests that other funding agencies could follow a similar pattern in designing their calls for grant proposals.

Research grants are a relevant setting for analyzing the impact of competition on participating agents. We found that awarded researchers are not the only beneficiaries in a research grant competition since the mere act of being involved in the process is beneficial in many aspects. This finding might be extended to other competitive contexts such as start-ups running for venture capital funding or firms applying for calls of public procurement contracts. When looking for investors to sustain their business, entrepreneurs are asked to write demanding business plans and to have precise strategic planning. Hence, the efforts spent in performing these requirements could be useful to improve the business performance of the start-up regardless of the result of the funding decision of investors. Similarly, firms could boost the efficiency of their projects when working on meeting the requirements of a public procurement call and benefit whether they succeed in winning the contract or not.

To our knowledge, this study is the first to consider the effect of applying on the scientific outcomes of researchers. Our results proved to be robust across several estimation strategies suggesting a treatment effect of the application process. However, one might still claim that our estimation strategy does not entirely account for the selection bias. We believe that this concern could justify a reduction of the magnitude of our results but not their direction. Moreover, we are aware that SINERGIA is a peculiar grant which might attract highly experienced scientists with interdisciplinary profiles thus limiting the scope of our findings. However, this paper intends to open the way for a better consideration of the effect of participating in competitions and future

studies could further investigate the mechanisms that lead applicants and non-applicants with similar profiles, to perform differently.

Acknowledgments

We are indebted to Gianni De Fraja, Dominique Foray, Adam Jaffe, Jacques Mairesse, and Reinhilde Veugelers for their valuable advice. We also thank Philippe Aghion, Guido Buenstorf, Gaétan de Rassenfosse, Enrico Forti, Martin Fransman, Rachel Griffith, Francesco Lissoni, Valerio Sterzi, and the seminar participants at the University of Bordeaux, University of Lausanne, University of Maastricht, University of Nice Sophia Antipolis, the conferences participants in AFSE 2017 in Nice, DRUID17 at NYU, Barcelona GSE, SEM at the Massachusetts Institute of Technology (MIT), and Academy of Management 2017 and three anonymous reviewers for their comments on earlier versions of the manuscript. We finally thank the SNSF for sharing with us their data. Any error in the document is our own.

References

Aghion, Philippe, Nick Bloom, Richard Blundell, Rachel Griffith, and Peter Howitt. "Competition and innovation: An inverted-U relationship." *The Quarterly Journal of Economics* 120 (2) (2005), 701-728.

Angrist, Joshua D., and Jörn-Steffen Pischke, *Mostly harmless econometrics: An empiricist's companion*. (Princeton university press., 2009).

Arora, Ashish, and Alfonso Gambardella, "The impact of NSF support for basic research in economics," *Annales d'Economie et de Statistique* 79/80 (2005), 91–117.

Arrow, Kenneth. "Economic welfare and the allocation of resources for invention." *The rate and direction of inventive activity: Economic and social factors*. (Princeton University Press, 1962. 609-626).

Ayoubi, Charles, Michele Pezzoni, and Fabiana Visentin, "At the Origins of Learning: Absorbing Knowledge Flows from Within or Outside the Team?," *Journal of Economic Behavior and Organization* 134 (2017), 374–387.

Azoulay, Pierre, Graff Zivin, Joshua S., and Manso, Gustavo. "Incentives and creativity: evidence from the academic life sciences." *The RAND Journal of Economics*, vol. 42, no 3, (2011) p. 527-554.

Azoulay, Pierre, Toby Stuart, and Yanbo Wang, "Matthew: Effect or Fable?," *Management Science*. 60(1) (2014), 92–109.

Azoulay, Pierre, Joshua S. Graff Zivin, Danielle Li., and Bhaven N. Sampat, "Public R&D investments and private-sector patenting: evidence from NIH funding rules," *the National Bureau of Economic Research working papers* No. 20889 (2015).

Carayol, Nicolas, and Marianne Lanoë, "The Impact of Project-Based Funding in Science: Lessons from the ANR Experience" *mimeo*. Université Bordeaux IV (2017).

Cole, Jonathan R., and Harriet Zuckerman. "The productivity puzzle." *Advances in Motivation and Achievement. Women in Science. JAI Press, Greenwich, CT* (1984).

Defazio, Daniela, Andy Lockett, and Mike Wright, "Funding incentives, collaborative dynamics and scientific productivity: Evidence from the EU framework program," *Research Policy* 38(2) (2009), 293–305.

Etzkowitz, Henry, "Research Groups as "Quasi-firms": The Invention of the Entrepreneurial University," *Research Policy* 32(1) (2003): 109—21.

Fox, Mary Frank. "Publication productivity among scientists: A critical review." *Social studies of science* 13.2 (1983): 285-305.

Graves, Nicholas, Adrian G. Barnett, and Philip Clarke, "Funding grant proposals for scientific research: retrospective analysis of scores by members of grant review panel," *BMJ* 343 (2011), d4797.

Gush, Jason, Adam .B. Jaffe, Victoria Larsen, and Athene Laws, "The Effect of Public Funding on Research Output: the New Zealand Marsden Fund," *the National Bureau of Economic Research working papers* No. 21652 (2015).

Huber, George P., “Organizational learning: The contributing processes and the literatures” *Organization Science* 2 (1991), 88–115.

Ioannidis, John P. A., “More time for research: fund people not projects,” *Nature* 477 (2011), 529–531.

Jacob, Brian .A., and Lars Lefgren, “The impact of research grant funding on scientific productivity,” *Journal of Public Economics* 95 (2011), 1168–1177.

Jaffe, Adam B., “Building programme evaluation into the design of public research-support programmes,” *Oxford Review of Economic Policy* 18 (2002), 22–34.

Lane, Julia, and Stefano Bertuzzi, “Measuring the Results of Science Investments,” *Science* 331 (2011), 678–680.

Lee, Sooho, and Barry Bozeman, “The Impact of Research Collaboration on Scientific Productivity,” *Social Studies of Science* 35 (2005), 673–702.

Long, J. Scott. "Measures of sex differences in scientific productivity." *Social Forces* 71, no. 1, (1992). 159-178.

Mairesse, Jacques, and Michele Pezzoni. "Does Gender Affect Scientific Productivity?." *Revue économique* 66, no. 1 (2015), 65-113.

Mora, Ricardo, and Iliana Reggio. "Alternative diff-in-diffs estimators with several pretreatment periods." *Econometric Reviews* (2017): 1-22.

Ruben, Adam, “Another tenure-track scientist bites the dust,” *Science* (2017). doi: 10.1126/science.caredit.a1700056.

Stephan, Paula, "The economics of science." *Journal of Economic Literature* 34, no. 3, (1996) 1199-1235.

Stephan, Paula, "The Economics of Science: Funding for Research," *International Centre for Economic Research Working Paper* No. 12 (2010).

Uzzi, Brian, Satyam Mukherjee, Michael Stringer, and Ben Jones, "Atypical Combinations and Scientific Impact," *Science* 342 (2013), 468–472.

Wooldridge, Jeffrey M., *Introductory econometrics: A modern approach*. (Cengage Learning, 2012).

Appendix

Section A – Discipline definition

We assign each scientist to one or more disciplines according to the journal where she published during the five years preceding the application year. A researcher is assigned to a discipline if she published at least one article in one of the journals classified in that discipline. We retrieve the classification of journals using the journal classification of the Scopus Source List[18]. When excluding Humanities and Social Sciences, the journal classification table reports 20 disciplines. We exclude generalist journals and journals belonging to more than one discipline to avoid ambiguous classification of researchers. Table A1 reports the non-exclusive attribution of disciplines to the sample of applicants.

Table A1: List of disciplines of applicants (1,060 observations).

SCOPUS disciplines	Mean	Std. Dev.	Min	Max
Chemistry	0.14	0.34	0	1
Engineering	0.09	0.29	0	1
Materials Science	0.05	0.22	0	1
Physics and Astronomy	0.20	0.40	0	1
Agricultural and Biological Sciences	0.10	0.30	0	1
Biochemistry, Genetics and Molecular Biology	0.42	0.49	0	1
Environmental Science	0.07	0.25	0	1
Computer Science	0.08	0.28	0	1
Mathematics	0.02	0.15	0	1
Medicine	0.32	0.47	0	1
Health Professions	0.00	0.03	0	1
Nursing	0.00	0.04	0	1
Earth and Planetary Sciences	0.08	0.27	0	1
Energy	0.01	0.09	0	1
Pharmacology, Toxicology and Pharmaceutics	0.06	0.25	0	1
Chemical Engineering	0.02	0.13	0	1
Neuroscience	0.17	0.38	0	1
Veterinary	0.01	0.11	0	1
Immunology and Microbiology	0.11	0.32	0	1
Dentistry	0.01	0.08	0	1

The table reports mean, standard deviation, minimum and maximum for the 20 dummy variables used to identify the scientists' discipline(s). The attribution of a discipline to a scientist relies on the discipline classification of the journals where she publishes. The classification of the journals is obtained from the bibliometric database Scopus Source List (Elsevier). The attribution of a scientist to a discipline is not exclusive, a scientist might be classified in more than one discipline.

[18] https://www.scopus.com/sources

Section B – Dealing with endogeneity issues in evaluating the impact of applying for a grant

In this section, we consider the possible endogeneity concerns that could be raised in estimating the effects of applying for a grant. We thus propose an alternative estimation strategy using a two-stage least square (2SLS) approach. We find that the core results presented in the main text are robust to the choice of estimation strategy.

Better researchers might self-select themselves into applying since they believe they have higher chances to succeed in being funded. Having high scientific profiles, applicants might perform better than other scientists regardless of the involvement in the application process. To address this selection bias concern, we follow an instrumental variable approach in addition to the propensity score matching strategy described in paragraph 3.2 of the main text. Specifically, we identify two excluded instruments for the *Applicant* variable. As first instrument, we consider the attractiveness of the researcher's discipline for funding agencies (*Funds in the US*). As second instrument, we consider the presence, in the researcher co-authorship network, of a scientist who applied for SINERGIA in the previous years (*Network Applicant*).

The first excluded instrument, *Funds in the US*, is a continuous variable that measures the availability of funds in the scientist's discipline during the application year[19,20]. The researcher who belongs to a discipline characterized by an extensive availability of funds is expected to perceive higher chances of success for her application, and she is then incentivized to apply. Given this correlation between *Funds in the US* and the dummy *Applicant*, the instrument is strong. It is also a valid instrument since the perception of higher chances to be awarded is not expected to

[19] Data on yearly US funds by discipline are available at https://www.aaas.org/page/historical-rd-data

[20] The available funds are adjusted for the size of each discipline as represented by the number journals classified as belonging to the discipline in the Scopus Source List 2017 journal classification table. (See also Section A of the Appendix)

influence the researcher's outcomes. The funds considered are those available in the researcher's discipline in the US (and not in Switzerland) to avoid any possible correlation with the unobserved factors that determine the researchers' outcomes in our sample. Given that scientific trends are not constrained by a country's boundaries; we expect disciplines highly funded in the US to be also highly funded in Switzerland. We attribute a value of funding to each researcher-application year according to the discipline(s) where the research is active (see Section A for discipline attributions)[21].

The second excluded instrument, *Network Applicant*, is a dummy variable that equals one if at least one of the co-authors of the researcher or, one of the co-authors' co-authors, has applied for a SINERGIA grant in the five years preceding the application year[22]. We expect *Network Applicant* to be negatively correlated with the probability of applying for SINERGIA. The reason for this expected negative correlation is that the SINERGIA call requires an applicant to find partners for her application. An applicant usually looks for co-applicants in her professional network. If the researchers in her network have already applied, they are less likely to be part of another application regardless of the result of their application. If they were awarded, there is a weak incentive for them to apply again. If they were not awarded, they might be discouraged to apply again. In both cases, the pool of potential co-applicants of the focal researcher is reduced with a detrimental effect on her probability to apply because she has greater difficulties in finding valuable partners for the application. Hence, we expect *Network Applicant* to be a strong instrument. Concerning the validity of the instrument, as suggested by the homophily mechanism

[21] If the researcher is classified in more than one discipline, we consider the sum of the funds available for each discipline (i.e. the propensity to apply of a researcher conducting her research in medicine and engineering will be affected by the funds made available in both disciplines).
[22] For a potential applicant, we consider as application year the year when her matched applicant submitted her application.

of Cummings and Kiesler (2008)[23], the choice of collaboration partners is mainly based on sociological characteristics or geographical considerations. The excluded instrument *Network Applicant* is therefore not directly correlated with the unobserved ability of researchers driving her expected scientific outcomes. Table B1 reports the results of our estimation exercise.

We test for the presence of endogeneity in the estimations using a *Durbin-Wu-Hausman* test. The test does not reject the null hypothesis that the variables are exogenous. Given the results of the *Durbin-Wu-Hausman,* OLS appears as the most efficient estimation, and its estimations are consistent with our main results (Table B2).

Table B1: Regression results for the scientific outcomes comparing applicants to potential applicants. IV estimations (2SLS).

	(1) IV Log(Publication count)	(2) IV Log(Average IF)	(3) IV Log(Average citations)	(4) IV Co-applicant collaboration	(5) IV Log(Learning)	(6) First step Applicant
Applicant	0.69***	0.12	-0.31***	0.66***	0.46***	
	(0.11)	(0.082)	(0.10)	(0.14)	(0.070)	
Funds in the US						0.082***
						(0.029)
Network applicants						-0.35***
						(0.026)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2,120	2,120	2,120	2,120	2,120	2,120
R-squared	0.642	0.496	0.498	0.486	0.584	0.403
Durbin-Wu-Hausman	0.16	0.60	0.37	0.84	0.24	

The table reports an instrumental variable estimation of the effect of applying for a SINERGIA grant (*Applicant*). The controls include scientist's demographic and bibliometric characteristics. The five scientific outcomes considered are *Publication count, Average IF, Average citations, Co-applicant collaboration*, and *Learning*. Column 6 reports the first stage equation of the 2SLS models. The dependent variable in column 6 is the dummy variable *Applicant*. The two excluded instruments are the variables *Funds in the US* and *Network applicants*. The sample includes 2,120 observations, i.e. 1,060 Applicant-application pairs, and 1,060 Potential applicants. Robust standard errors are clustered around the application. In reporting the statistical significance of the coefficients, we apply the standard thresholds, namely *** p<0.01, ** p<0.05, * * p<0.1.

[23] Cummings, Jonathon N., Kiesler, Sara, "Who collaborates successfully? prior experience reduces collaboration barriers in distributed interdisciplinary research" *Proceedings of the 2008 ACM Conference on Computer Supported Cooperative Work. ACM*, pp. 437–446 (2008).

Table B2: Regression results for the scientific outcomes comparing applicants to potential applicants. OLS estimations.

	(1) OLS Log(Publication count)	(2) OLS Log(Average IF)	(3) OLS Log(Average citations)	(4) OLS Co-applicant collaboration	(5) OLS Log(Learning)
Applicant	0.53***	0.081***	-0.41***	0.50***	0.48***
	(0.035)	(0.031)	(0.033)	(0.047)	(0.026)
Controls	Yes	Yes	Yes	Yes	Yes
Observations	2,120	2,120	2,120	2,120	2,120
R-squared	0.645	0.496	0.499	0.489	0.584

The table reports an OLS estimation of the effect of applying to the SINERGIA grant (*Applicant*). The controls include scientist's demographic and bibliometric characteristics. The five scientific outcomes considered are *Publication count, Average IF, Average citations, Co-applicant collaboration,* and *Learning*. The sample includes 2,120 observations, 1,060 Applicant-application pairs, and 1,060 Potential applicants. Robust standard errors are clustered around the application. In reporting the statistical significance of the coefficients, we apply the standard thresholds, namely *** p<0.01, ** p<0.05, * * p<0.1.

Section C – Entering new research domains

As discussed in Section 6 of the main text, we measure the entry in new research domains relying on the knowledge mobilized by a researcher in her scientific publications. Precisely, we measure the cognitive distance between the researcher's learning -i.e. the new journals cited after the application date by a scientist- and her initial stock of knowledge –i.e. the journals cited in the articles published before the application year. To do so, we build on Ayoubi et al. (2017) to construct a journal distance matrix measuring the distance between all existing pairs of journals. This section reports the details of the construction of the matrix and the steps leading to the creation of the variable *Journal distance* used as dependent variable in Table 8 in the main text.

The journal distance matrix

We build the journal distance matrix based on the assumption that the closer two journals are in terms of scientific content the more likely they are to be co-cited in scientific publications. Exploiting this assumption, we calculate the closeness of two journals as the ratio of the number of publications in which they are co-cited to the number of publications in which the least cited journal of the two is cited. Then, the distance reported in our journal distance matrix is the inverse of the computed closeness. Specifically, we derive the distance separating two journals *i* and *j* as the inverted ratio between the number of publications in which *i* and *j* are cited together and the minimum number of publications where *i* or *j* are cited depending on which one is the least cited. This computed distance measure ranges from 1 to infinity. For the infinite values, namely when *i* and *j* are never cited together in a publication, for computational purposes, we consider as distance the maximum non-infinite distance found in the matrix. Equation C1 shows the calculation of the journal distance for two journals *i* and *j*.

$$D(i,j) = \frac{1}{\frac{\text{\#pubs where i and j are co-cited}}{\min(\text{\#pubs where i is cited, \#pubs where j is cited})}}$$

(Equation C1)

We calculate the journal distance matrix based on the publications of Swiss scientists active in the top 12 universities in Switzerland (excluding humanities and social sciences) in the period 2003-2015, for a total of more than 200,000 articles. The journal distance matrix includes 3,013 journals. The minimum distance value is 1, while the maximum distance is 5,628. In the computed matrix, 65% of the journal pairs never appear together in the reference list of an article. For those pairs, we fix the distance value at the maximum. For the sake of illustration, we consider an extraction from the journal distance matrix including two journals in physics, i.e. *Physical Review Letters* and *Annals of Physics*, and a journal in molecular biology, i.e. *The EMBO journal*. Table C1 shows the distance between the three journals in the journal distance matrix extraction.

Table C1: Journal distance matrix extraction.

	Annals of Physics	*The EMBO journal*	*Physical Review Letters*
Annals of Physics	-	119	2
The EMBO journal	119	-	66
Physical Review Letters	2	66	-

The table shows an extraction of the journal distance matrix resulting from the calculation of the distance between each pair of journals. The distance calculated between two journals is based on the assumption that the closer two journals are in terms of scientific content the more likely they are to be co-cited in scientific publications.

As expected, the two journals in physics are close meaning that they are frequently co-cited in the references of published articles (D(*Physical Review Letters, Annals of Physics*)=2). These two journals in physics are far from the journal in biology (*The EMBO journal*) with a distance value of respectively D(*The EMBO journal, Annals of Physics*)=119 and D(*The EMBO journal, Physical Review Letters*)=66.

We use the journal distance matrix to calculate the average cognitive distance between the scientist's stock of knowledge before the application year and the knowledge she acquires after the application year. To calculate this cognitive distance, we proceed in four steps. First, we extract both the list of distinct journals cited by the scientist before the application and the list of the new journals cited after the application. Second, we construct the pairs resulting from all the possible combinations of distinct journals in the two lists of the first step. Third, we attribute to each journal pair its corresponding distance value in the journal distance matrix. Finally, we create the variable *Journal distance* as the average of the distances retrieved in step three. The variable *Journal distance*, in logarithmic form, is the dependent variable of the regression shown in Table 8 in the main text.