Stimulating Collaborations: Evidence from a Research Cluster Policy

Nicolas Carayol[*], Emeric Henry[†], and Marianne Lanoë[‡§]

November 19, 2022

Abstract

The production of knowledge relies on interactions and collaborations between researchers. However, we do not know to what extent policies may stimulate these interactions. In this paper we shed light on this question, and show how a public "research cluster" policy, which funds local communities of researchers working on a common theme, affects the organization of research within these clusters and the production of its members. Using data from a large-scale financing program in France, and relying on an identification strategy based on grades awarded by reviewers, we show that members of financed clusters increase their research collaborations with other members of the cluster by up to 30%, compared to researchers involved in nonselected proposals. Paradoxically, those who benefit the most from the program are those who were not at the core of the research topic; these researchers significantly increase their links

[*]Univ. Bordeaux, CNRS, BSE, UMR 6060, F-33600 Pessac, France. Email: nicolas.carayol@u-bordeaux.fr

[†]Department of Economics, Sciences Po Paris, 28 rue des Saint Peres, 75007 Paris (France)

[‡]Univ. Bordeaux, CNRS, BSE, UMR 6060, F-33600 Pessac, France, and Agence Nationale de la Recherche. Email: marianne.lanoe@agencerecherche.fr

[§]We are extremely grateful to the staff at the ANR, in particular Arnaud Torres, Emmanuelle Simon and Yves Lecointe who gave us access to the LabEx data and helped us collect it. We thank Pierre Gouedard for his invaluable help in collecting the data. We thank Matt Jackson, Joan Monras and participants in seminars and conferences for their comments. This research benefited from the financial support of the ANR (ANR-15-CE26-0005) and Bordeaux IdEx (ANR-10-IDEX-03-02).

with core members and move their research focus closer to the cluster's theme as a consequence of the policy.

1 Introduction

The production of knowledge, a key fuel for economic growth (Romer, 1990; Aghion and Howitt, 1992), critically depends on peer effects.[1] According to Borjas and Doran (2015), peers matter along three dimensions of proximity: in the collaboration space (as coauthors), in the geographical space (as researchers working in the same location) or in the idea space (as researchers working on the same topic). Although it remains essentially unknown whether policies can be designed to stimulate interactions among peers, we have recently witnessed policy initiatives explicitly aiming to do so by financing academic "research clusters", i.e. networks of researchers working in the same location (geographical space), on a common theme (ideas space) with the goal of encouraging new interactions (collaboration space).[2] In this paper, using a public policy experiment in France where a national contest was run to select and fund research clusters, we study how the policy stimulates collaborations, affects the nature of academic production and how researchers are differentially affected.

Exploiting an identification strategy based on grades given by reviewers, we show that the program led to a significant restructuring of the collaboration network of researchers. Members of selected academic clusters, compared to those working in clusters that were not chosen but received similar grades, increase their collaborations with other cluster members by up to 30% several years after the financing was obtained. They not only strengthen existing teams, but also initiate new ones. Furthermore, and seemingly paradoxically, the

[1] The literature on peer effects in the production of knowledge includes numerous papers such as Azoulay, Graff Zivin and Wang (2010); Waldinger (2012); Borjas and Doran (2012); Oettl (2012); Borjas and Doran (2015); Jaravel, Petkova and Bell (2018).

[2] The Exzellenzinitiative in Germany, the "Severo Ochoa" Centers of Excellence in Spain, the Centers of Excellence in the Nordic countries (descriptive evidence in Möller, Schmidt and Hornbostel (2016) and Langfeldt et al. (2015)) or the Initiative d'Excellence in France. Universities also increasingly divert funds from traditional discipline based funding to invest in specific themes. There are numerous instances of clusters (or centers) of excellence created recently within (or sometimes across) universities such as the University of British Columbia, Stanford University, MIT and the University of Cambridge.

researchers who are most affected by the policy are those not included in the bibliography of the proposal (whom we call "periphery members"), in other words those who were not working on the cluster research topic prior to the treatment. These researchers significantly increase their collaborations with "core members" (those cited in the bibliography) and move their research agendas closer to the cluster's theme.

The policy experiment we study is a large-scale research funding initiative in France, called the LabEx program. In two successive calls in 2010 and 2011, 436 research cluster proposals were submitted, of which 171 were funded. The average allocation was 8.8 million euros over 10 years. These clusters bring together researchers from different research units, not necessarily from the same institution, to work on a common theme. Our research design exploits different key pieces of information. First, from the agency running the program, we obtained the grades given by reviewers for all proposals, both those accepted and those rejected. Second, we were given access to some information contained in each project proposal from which we extracted the list of associated research units and the names of authors cited in the project's bibliography. This data allows us to define members of these clusters and, exploiting the bibliography, to determine in a very precise way the core and periphery members.

Our main analysis is a difference in differences specification, in which we restrict the sample to researchers who were part of only one cluster proposal.[3] We thus compare post financing members of financed clusters to research involved in nonfinanced proposals. We include time and researcher fixed effects to account for differences in levels between the two groups.[4] To increase the validity of the comparison, we restrict our analysis to researchers who received similar grades on their proposal. Specifically, we restrict to a range of grades where the probability of selection hovers between 20% and 80%. Given this restriction, the selection or rejection of their unique proposal can be considered essentially random, an

[3]We explain in details in Section 2.2.3 how we deal with multiple memberships and show that researchers involved in several proposals are similar to those participating only in one.

[4]We also include gender $\times$ year, age in 2010 $\times$ year, and discipline $\times$ year fixed effects to account for differential gender, discipline and cohort effects over time.

assumption we will support by examining pretrends.

Using our main identification strategy, we first show a very large effect of the policy on the organization of research. On average, the number of collaborations with coauthors from the same cluster increases by 16% for members of funded clusters due to the policy, an effect that ramps up to 30% several years after the policy was launched. The effect is mostly present for collaborations involving core members, particularly those bringing together periphery and core members. Importantly, we show the absence of any significant pretrends, thus supporting our identification strategy.

We also examine how the policy affected the production of science. In terms of standard productivity measures, such as the number of publications adjusted for quality, we do not find a clear impact of the policy, although the estimates are noisy. This type of cluster policy may affect the productivity of research for a given theme, but may also induce movements in the space of ideas. We explore this second possibility next. The research theme of each cluster is characterized using the keywords listed in the papers cited in the reference list of each cluster. Comparing the keywords used each year by cluster members to those characterizing their cluster, we show that periphery members move significantly closer to the cluster theme post treatment. Specifically, they publish more papers using at least one of those keywords, an effect in the order of a 22% increase. The policy induced periphery members, not only to build connections with core members, but also, and probably as a consequence, to move their research closer to the cluster's theme.

We conclude the study by discussing potential mechanisms that drive our main results on collaborations and justify why these newly formed collaborations, if productive, had not been formed previously. We base our discussion on a small survey of 12 managers of funded clusters that we conducted. All the respondents stated that they witnessed either the creation or strengthening of links between group members. They also mentioned that common seminars were created involving several founding members of the research cluster, that cosupervisions of PhD students were established and that students typically shared

common facilities. The first mechanism that we argue can explain our results, and that we formalize in a simple theoretical model, is one based on public good provision. The creation of the cluster decreased the cost of providing the public goods mentioned in the survey (e.g., seminars, cosupervision of PhDs, training). These public goods, in turn, increased the value of internal collaborations. This mechanism can lead to the observed restructuring of the network of collaborations and can also explain why periphery members, who are the beneficiaries of these public goods, have more to gain from the cluster policy, than core members, who are on the contrary providers. The second mechanism is based on the governance of these clusters. The newly formed entities had to create their own rules since no explicit guideline was given by the funding agency. Many chose to set up internal calls for funding where one of the requirements was to have members of at least two of the founding labs involved in the project. This requirement would naturally lead to an increase in collaborations.

The economics of science has recently focused on research funding and on the organization of research (Bryan and Williams, 2021). On the first dimension, a number of papers study the design and impact of research funding programs.[5] Some studies focus on the *ex ante* stage and examine how accurate reviewers are in predicting project outcomes (Li and Agha, 2015; Park, Lee and Kim, 2015; Fang, Bowen and Casadevall, 2016), or how biased evaluations are with respect to the characteristics of the PIs, the proposals or the reviewers (Ginther et al., 2011; Boudreau et al., 2016; Li, 2017; Banal-Estañol, Macho-Stadler and Pérez-Castrillo, 2019). *Ex post* studies examine the impact of fund allocation on the selected proposals, typically considering productivity as the variable of interest.[6]

A number of papers focus on the organization of academic research. In particular, several papers (Azoulay, Graff Zivin and Wang, 2010; Oettl, 2012; Jaravel, Petkova and Bell, 2018) exploit the unexpected deaths of scientists to estimate the causal effect on the productivity

[5]See Azoulay and Li (2020) for a survey and Azoulay, Graff Zivin and Manso (2011), Jacob and Lefgren (2011) Carayol and Lanoe (2019), Defazio, Lockett and Wright (2009)

[6]Jacob and Lefgren (2011) find a 7% increase in citations following NIH grants, Carayol and Lanoe (2019) in the French context also find positive effects on productivity.

of their coauthors or collaborators.[7] Others use historical shocks affecting departures or arrivals of researchers to quantify the impact on others. Waldinger (2012) shows that the scientists whose departments suffered losses during the period from 1925 to 1938 did not publish less or worse than other scientists. Borjas and Doran (2012) examines the impact of unanticipated mobility on the destination side. The authors show a negative effect of the influx of Soviet Union mathematicians in the US after the collapse of the Iron Curtain on the productivity of American mathematicians, due to competition for scarce resources, but find no effect on overall productivity.[8]

Our work lies at the intersection of these two strands of the literature. It contributes to the literature on science policy as it is the first paper to empirically assess the causal impact of funding research clusters rather than individuals or small teams. It also contributes to the literature on the organization of science in showing how funding of science affects the organization of academic research. In particular, we demonstrate that funding research clusters increases collaborations within the group. In addition, we also find interesting conjugated effects on the direction of research, as periphery members significantly reorient their research toward the theme of the research cluster proposal. Therefore, research cluster funding turns out to be a particularly efficient policy design to implement thematic movement toward promising research areas, whereas it has been highlighted that it is very costly to achieve the same outcome through a more standard funding design (Myers, 2020).

The rest of the paper is organized as follows. In Section 2, we present institutional details on the research cluster policy whose impact we investigate and the data we use. Section 3 gives evidence on the selection procedure and presents the identification strategy. Section 4 presents our main results on collaborations and on the production of science.

[7] Azoulay, Graff Zivin and Wang (2010) find a strong effect following the death of star scientists, while Oettl (2012) qualifies this result by showing that the effect is restricted to helpful scientists (i.e. those acknowledged in several papers per year). Jaravel, Petkova and Bell (2018), using patent data, show that the effect is not restricted to stars.

[8] See also Iaria, Schwarz and Waldinger (2018); Bosquet et al. (2022) and the literature on teams in science. Wuchty, Jones and Uzzi (2007), using several decades of data on publications, show that research is increasingly reliant on teamwork across fields and, furthermore, that the knowledge produced by teams is more likely to create very high impact research.

Section 5 discusses the mechanisms driving our main results and policy implications, building on interviews and a theoretical framework. Section 6 concludes.

2 Institutional Setting and Data

2.1 Institutional Setting

Based on a bipartisan report written by two former prime ministers, French president, Nicolas Sarkozy, announced in 2009 a large-scale investment plan for research and productivity. One essential element of this plan was the "Laboratoires d'Excellence" (LabEx) program. It aims at financing consortia of research units that plan to work on a common theme (which we refer to as research clusters). The stated goal of the program was to favor the emergence of ambitious scientific projects and to reinforce the visibility and strengths of the best labs. There was no explicit goal of fostering collaborations within the cluster.

The program was run on a bottom-up and fully competitive basis at the national level by the Agence Nationale de la Recherche (ANR).[9] A first call for research cluster proposals was issued in 2010. Each application involved several research units, the elementary blocks for the organization of research in France, with one coordinator in charge. The 241 applications were sent to external reviewers, and the independent international committee selected 100 winning proposals that were announced on March 25, 2011. In response to the second call for proposals issued in October 2011, 195 proposals were submitted (including 55 resubmissions from the first stage, described in more details in the Supplementary Appendix), of which 71 were funded.[10]

The funding was granted for an 8-year period, with an average allocation of 8.8 million euros, ranging from 2 to 30 million. In many cases, this amount allowed the clusters to

[9]The ANR was created in 2005 to perform grant-based research funding. Its organization has been redesigned to also administer this program.

[10]A large part of the prefinancing of the first wave was paid between July and November 2011, whereas the second wave was paid between May and August 2012, after which funds are transferred on a yearly basis. Figure A.1 in the Supplementary Appendix presents a map of the accepted clusters.

raise further funds. Each cluster is organized in a specific way, but most have a director, an executive and a scientific committee. Some also have a steering committee. In most cases, leading scholars were involved in top management[11] and the management burden is relatively low with respect to the national standards.

2.2 Data

2.2.1 Research Clusters

The ANR shared with us all the application files they received for the LabEx program, including those that were not selected.[12] All files include the name of the coordinator, the name and identifying codes of the partner research units, the amounts requested and the funding decision. It also includes a summary of the project.[13] In addition each file contains a bibliography from which we extract the names of all authors and the keywords used in these papers.

The ANR also provided us with an additional piece of information, essential for our identification strategy, namely the grades awarded by the referees to each proposal. External referees graded proposals on seven criteria: the quality of the teams and facilities, the relevance of the research project goals, the potential in terms of innovation and impact, involvement in training (especially master's degrees and PhDs), organization and management, institutional strategy (universities and research institutes), and project/means adequacy and ability to generate resources. We obtained the grades separately by criterion.

[11]For instance, Jean Tirole is the President and Chairman of the Executive Committee of the IAST (Institute of Advanced Studies in Toulouse) research cluster, dedicated to the interactions between economics and other disciplines.

[12]The ANR shared with us 200 of the 241 files for the 2010 call, removing the proposals that received the lowest grade, and were not close to being selected. For the 2011 call, they shared all the files with us.

[13]For confidentiality concerns, we were not given access to the full text of the proposals.

2.2.2 Cluster Membership

The cluster proposals that we obtained from the ANR did not contain a list of participating researchers (except for the coordinator), but included a list of participating research units. We define research cluster membership, for both the selected and nonselected proposals, as being a professor or researcher member of a research unit listed as a founding partner in the proposal file described above. This is a broad definition since not all members of the associated units necessarily effectively participate in the activities of the research cluster.

Given this definition, we had to recover the list of all members of the research units listed in the cluster application. We used a country-wide roster that contains approximately 96% of all professors and academic researchers who have been employed in academia in France since 2005. This unique dataset (data sources are described in the Appendix) has been built using a variety of official sources. As it offers information on the research unit to which the individual is affiliated, it was possible to link each person to a cluster proposal. We find, given our definition of membership, that approximately 49 thousand researchers were members of a cluster proposal, which is approximately half of the initial population.

In addition, we use data extracted from the project bibliography to further refine the membership definition. For each research cluster, we identify two subgroups:

- The **core members** are members of the research cluster (i.e. members of one of the research units listed as partners) who also appear as authors of articles listed in the bibliography of the proposal.

- The **periphery members** are members of the research cluster whose work is not referenced in the bibliography.

This should be a relatively precise measure of core and periphery members, since it was in the interest of those writing the cluster proposal to include all the relevant papers in the field. There are two potential sources of error. On the one hand, we may include researchers cited in the bibliography who are no longer working in the field or even active in research. On the

other hand, we may miss researchers not cited because they have just started their career. However, this appears unlikely since descriptive statistics in the Supplementary Appendix (Table A.2) show that core and periphery members are relatively similar in age.

2.2.3 Multiple Memberships

Clusters typically involve many research units as shown in Table 1 (15 on average, most have less than 10, but some have significantly more). This implies that, with the definition of cluster membership introduced above that assigns membership to all researchers in a unit, many individuals are members of several cluster proposals (64% of researchers are in more than one cluster proposal). Such a form of multiple membership has implications for the study design that are extensively discussed in the next section.

There likely exists a second reason for this high figure: there was a tendency by those writing the proposals to inflate the number of units involved in the hope of increasing their odds of success. A simple way to appreciate the involvement of a research unit in a research cluster is to look at the proportion of its members who appear in the bibliography of the proposal. Figure A.2 of the Supplementary Appendix plots, for all unit $\times$ cluster pairs, the distribution of the proportion of researchers in the unit who appear in the bibliography of the proposal. We observe that 10% of units have no member appearing in the bibliography, which suggests their involvement is quite limited, at least at the time the team was assembled.[14]

2.3 Variables

To capture the effect of the policy on the organization, productivity and direction of research, we collect bibliometric information on all cluster members. For that purpose, we first match all cluster member names to the authors of scientific articles (on the basis of surname and first name initials) in the Clarivate Web of Science (in-house XML datafiles and online access), which centralizes all the documents published in the main scientific journals. We retrieve

[14]This gives rise to a specification (specification (4) in our main tables) testing the robustness of our results as explained in Section 3.

more than 10 million documents published until 2019 that need to be filtered to keep only those that have been authored by professors and researchers in the research clusters we study, excluding homonyms from around the world. We do so using all available information we have on individual profiles through a "seed and expand" methodology (Reijnhoudt et al., 2014) and up-to-date machine learning algorithms.[15]

We focus our analysis on two key outcome variables, measuring the intensity of collaborations within the clusters and the productivity of researchers. The first variable, *Links*, measures the number of links, i.e. coauthors, an individual has within the cluster. We argue in the theoretical framework discussed in Section 5 that co-authorships within the cluster should increase in financed clusters after the start of the program. As a robustness measure, we also explore the effects on other variables measuring the organization of research: the number of papers with at least one other author from the cluster (*CollaPubs*), the corresponding measure of the number of papers without coauthors from the cluster (*ExternalPubs*) and the number of coauthorships within the cluster that had not occurred before (*NewLinks*). We also separately consider links with core and periphery members of the cluster (*LinksCore* and *LinksPeriph*).

To measure the productivity of researchers we mainly focus on the variable *AIF* which counts the number of publications weighting each paper by the journal impact factor. It is a measure of research outcomes adjusted for quality. The journal impact factor is a standard measure to assess a journals' quality. We calculate our own measure of the impact factor defined as the average number of citations that a paper published in year t receives from papers published in $[t; t+2]$. We chose this measure as it is more qualitative than just

[15]In a first step (seed), the algorithm validates articles by imposing strong and reliable conditions, particularly on the scientific field and host institution that must be fully consistent with the information we have for each person. The "expand" stage is in fact composed of a series of loops in which information on validated articles is used to make decisions on articles that pass only some of the conditions imposed in the "seed" stage. For instance, if it turns out that a candidate paper has the same coauthors, cites the same references, or uses the same keywords as validated articles, this increases the conditional probability of a correct match. We use machine learning algorithms that are trained on a subset of French professors and researchers who have created an ORCID identifier and are thus likely to have carefully selected their own publications. From end to end, this filtering process is controlled for and fine tuned to improve efficiency in terms of precision and recall.

counting the number of articles published and it is less noisy than weighting papers by 3-year forward citations. We however also consider effects on those two variables (*Pubs* and *Cites*) as robustness exercises.

As the cluster policy aims at fostering scientific excellence, we also build indicators focusing on top "quality" scientific outcomes. For each researcher, we count the yearly number of papers that are among the top 5% and top 10% most cited papers in the research field (*Top5* and *Top10*).[16]

We also want to appreciate if the policy affects the direction of research. To explore this potential impact, we exploit the bibliography of each cluster, which was already used to define core and periphery members. We extract the papers cited in the bibliography of each cluster and use the keywords obtained from Web of Science to define its research theme. A keyword may have different meanings in different fields and we therefore also make use of the information on the research field of each paper. Accordingly, a universe of pairs of keywords – subject categories associated with each cluster is built to characterize its research theme, which we refer to as the cluster *universe*. Altogether we obtain approximately 28 thousand triples cluster $\times$ keyword $\times$ subject category. This allows us to count the yearly published number of papers using at least one keyword $\times$ subject category of the cluster universe authored by each of its members (*ClusterTheme*).

Finally, we collect a number of control variables. At the researcher level we observe age, gender and scientific field. Six broad fields are defined: biology, chemistry, physics, engineering, mathematics and social sciences and humanities. At the cluster level, we also construct a number of variables beyond the grades discussed above. The field of specialization of the cluster is defined as the most frequent field among its members. We also count the number of involved research units and the proportion of core members.

[16]Formally, top cited over a three-year period and among papers of the same type (research article, letter, review) in the same Web of Science "subject category" that divides science in 252 fields.

3 Identification Strategy and the Selection Process

In this section we set the stage by first empirically examining how proposals were selected based on grades and presenting the average characteristics of the research clusters and researchers involved. This selection process is the basis of our identification strategy described in Section 3.2.

3.1 Research Clusters and Selection Process

We present in Table 1 the average characteristics of the funded research clusters (Column (1)) and compare with those that were rejected in Column (2) (Column (3) presents the comparison of means test). On average we observe 275 researchers in each research cluster with a relatively large variance. Of those, 31% are core members. The selected clusters tend to be larger and more productive than the rejected ones. A feature important for us in the rest of the analysis, is that the proportion of core members is very similar across selected and nonselected clusters, 31% versus 32%.

As described in Section 2.2, we obtained the grades awarded to each proposal and, therefore, provide some evidence on the selection process. Table 1, also compares the grades of the selected projects to the grades of those rejected, distinguishing the different components of the grade. Reassuringly, grades are significantly higher for those selected, and this is true for all the components of the overall grade. The difference between the grades of those selected and rejected is particularly large for the grade on the quality of the team (criterion 1), goal of the project (criterion 2) and potential for research output (criterion 3).

As described in Section 2.1, the final decision of which academic clusters to select did not follow a cutoff rule. We plot in Figure 1 Panel A, the probability of being accepted as a function of the total grade, distinguishing between the 2010 (on the left) and 2011 contests (on the right). Consistently, the probability of acceptance increases with the grades. However there is a range of grades where the probability of acceptance hovers between 20% and 80%.

We later restrict the analysis to this intermediate range of grades, to compare projects with similar latent characteristics. We then plot in Figure 1 Panel B, the probability of being accepted as a function of the grade on criterion 3, which corresponds to the research potential of the project. We see that receiving a grade of 4 out of 5 on this dimension, results in an approximately 50% chance of being selected in the 2011 contest. This will also be used to construct an alternative identification strategy.

3.2 Identification

We estimate the following difference-in-difference specification in our main tables:

$$y_{ict} = \mu \text{ Treatment}_{ic} \cdot \text{Post}_{ct} + \beta X_{ict} + \gamma_i + \eta_t + \epsilon_{ict}, \tag{1}$$

where y_{ict} is the outcome variable of interest, Treatment_{ic} identifies whether individual i is a member of a research cluster c that was funded, and Post_{ct} identifies the post policy period. This specification includes individual (γ_i) and year (η_t) fixed effects, which capture possible differences in levels. As shown in the descriptive statistics of Table 2, researchers in funded projects tend to be more productive than those in non funded ones, in all dimensions of productivity (publications, citations, coauthorships). Finally, X_{ict} are time varying fixed effects including gender $\times$ year, cohort $\times$ year, and discipline $\times$ year fixed effects that account for possible differential gender, discipline and cohort effects over time.

In this specification (1), the unit of analysis is the individual researcher $\times$ cluster $\times$ year level, although the financing is at the research cluster level. We adopt this approach because of the issue of multiple membership discussed in section 2.2.3. In a given research cluster, some members might also belong to other cluster proposals, whereas others would participate only in this specific structure. This naturally poses challenges for the identification of the effects of cluster policy as some individuals may be members of both treated and non treated clusters. Moreover, some researchers may be members of several treated clusters raising the

issue of how these multiple treatments add up. Thus, any post-treatment outcome for a given cluster may typically include the outcomes of treated researchers, potentially several times, even though the cluster itself has not been funded. We therefore refrain from undertaking the impact analysis at the cluster level and instead undertake the study at the individual scientist – research cluster pair. This allows us to adopt a conservative approach restricting our analysis to researchers involved in one single cluster proposal (which can be ultimately accepted or rejected). To ensure the external validity of our exercise, we show in Table A.3 that those researchers are very comparable to those involved in two or three clusters in terms of pre-treatment variables (we take the average value of the variable in years before 2011), in particular on variables reflecting the number of collaborations which is our main focus.[17]

As we restrict the analysis to the individuals who are part of a single cluster, the analysis in Equation 1 boils down to an individual researcher $\times$ year level study. We however keep the cluster subscript for clarity. Besides, given that observations may be correlated within each research cluster, we always cluster standard errors at that level (which nests the clustering at the individual level since the observations of any individual are limited to one single research cluster).

The key assumption to ensure the validity of the DID approach is the parallel trend assumption. This requires that in the absence of treatment, the difference between the treatment and control groups is constant over time. However, without further restrictions, there are some reasons for these trends to potentially differ. For instance, if the committee is good at selecting more promising projects with better potential, the results would be biased upward. In contrast, if the committee overweighs past achievements that might be negatively correlated with future trends, in other words, selects research clusters that just reached their peak, the bias would go in the opposite direction.

To overcome these challenges to identification, our main strategy exploits the grades as-

[17]Some cluster proposals were not accepted in the first contest and resubmitted in the second one. We treat them as separate submissions since they often involve a change in both team and focus. These projects are therefore not included in our data. We discuss resubmissions in more details in an online appendix.

signed by reviewers to the different projects. Specifically, we restrict our analysis to clusters that received a grade in a specific range. We chose this range as the set of grades that gave a probability of approval between 20% and 80%.[18] We thus restrict our sample to projects having ex ante similar potential, interpreting the final selection as reflecting orthogonal factors such as geographical coverage of the territory. Table A.1 in the supplementary appendix shows that treated and non treated samples before treatment are very similar. Besides, we verify graphically for our main variables of interest that when we impose this restriction, there are no significant pretrends between the control and treatment groups. This is not a statistical test of the parallel trend assumption, but rather evidence compatible with that assumption.

In our main tables, we consider three additional specifications. First, we consider a specification (column (2) of our main tables) that replaces in specification (1) the individual fixed effect by fixed effects for grade point on criterion 3 (potential of the project) $\times$ discipline $\times$ application year. This specification exploits differences between funded and non funded clusters, in the same discipline, applying in the same year and receiving the same grade.

Second, we consider a specification that addresses the issue of possible spillovers from funded to non-funded clusters that could constitute a potential threat to our empirical strategy (column (3) of our main tables). Indeed, a treated unit might be located close (geographically) to a nontreated cluster in the same field. This could lead us to either over- or underestimate the treatment effect depending on the sign of the spillovers.[19] To account for this possibility, we consider specification (1) but restrict to research clusters for which there is no other research cluster in their field within a 10 km radius.

Finally, the last specification we use when presenting our main results (column (4) of our main tables) addresses the concern highlighted at the end of Section 3, the fact that

[18]The specific condition is that the cluster received an overall grade between 26 and 32 for the 2010 contest and between 30 and 32 for the 2011 contest.

[19]If the treated unit produces local public goods (workshops, invited researchers) that benefit the nontreated cluster, we would underestimate the effect. In contrast, if the financed LabEx attracts activity and partnerships far from the nonfinanced cluster, the spillover would go in the opposite direction.

certain units might have been strategically added to the proposals without being close to the cluster's theme. We estimate specification (1) but remove from the sample the researchers $\times$ cluster pairs when none of the researchers in the associated research unit appear in the bibliography of the proposal.[20]

In the Supplementary Appendix we consider some additional robustness checks. We use a strategy that fully exploits a particular component of the grade measuring research potential (criterion 3), restricting to cluster proposals that received a grade of 4 out of 5 on that criterion. These projects were judged by the reviewers to have the same potential future trends, but some were financed for reasons possibly orthogonal to research output, such as the potential to generate training (criterion 4) or the strategy of the supervising institutions (criterion 6).[21] We also conduct three additional robustness exercises. First, for our main tables we apply a stronger restriction on grades.[22] We also use Coarsened Exact Matching (CEM) techniques (Iacus, King and Porro, 2012) to match treatment and control observations based on pretreatment characteristics. The approach essentially compares similar scientists (same age and research profile) in similar research clusters (same field and grade), but some are treated and others are not. Last, we use assume that the treatment may be dynamic with respect to the time to treatment. The ATT is here first estimated for each year and each time to treatment and then aggregated (Callaway and Sant'Anna, 2021). See below our description of event studies in Subsection 4.1.[23]

4 The Impact of Research Clusters Funding

The research cluster funding policy we study targeted relatively large groups of individuals with the ambition of affecting the interactions between them. This is what makes this funding

[20]This redefines the whole sample we use since it affects the number of clusters in which individuals participate.

[21]Note that the projects receiving a grade of 4 on this criterion had roughly a 50% chance of being selected.

[22]Specifically, for this stronger restriction, we require the grade of the unique proposal to be between 28 and 31 for 2010 and equal to 31 in 2011, which guarantees a probability of selection between 40 and 60 %.

[23]A proper description of the assumed data generation process here is similar to that of Equation 2 below.

instrument potentially different from an aggregation of individual grants. We therefore start by investigating whether the program reshaped research links.

4.1 Organization of the Research Network

To illustrate the potential effects on collaborations and to preview our empirical results, we start by presenting in Figures 2 and 3 an illustration of two research clusters, one funded and the other rejected. We present the graph of the collaboration networks of these two cluster proposals (where a link is defined as having a copublication over the period) separately before and after 2012, i.e. the first treatment year for the funded cluster. For representation purposes, we also restrict the representation to nodes corresponding to individuals involved in a unique cluster proposal. The two research clusters we present received similar grades and had a similar number of nodes (58 and 60 for the treated and the nontreated respectively) and similar numbers of links before treatment (69 and 74 respectively). We represent the periphery members in blue and the core members in green. The first striking feature visible in Figure 2 is that, for the funded cluster, the network becomes much denser in the period after treatment, while changes are more limited for the rejected cluster in Figure 3. The second salient feature is that the periphery members represented in blue significantly increase their connectivity in the funded cluster with core members.

We show in Table 3 that the first feature differentiating the evolution of the two research clusters discussed above applies much more generally.[24] The first column in Table 3 reports the results of our main specification, which we restrict to researchers involved in a unique proposal that obtained similar overall grades. We see that there is a large and significant effect on the number of collaborations within the financed clusters after the start of the program. Internal links increase by nearly 16% in response to the cluster policy.

In Figure 4 we present the results graphically. Specifically we assume the following data

[24]We address the second feature later in this section.

generating process:

$$y_{ict} = \sum_{\tau, t_c} \mu_{\tau t_c} \text{ Treatment}_{ic} \cdot \mathrm{D}(t, t_c, \tau) + \beta X_{ict} + \gamma_i + \eta_t + \epsilon_{ict}. \tag{2}$$

Function $\mathrm{D}(\cdot)$ is defined as $\mathrm{D}(t, t_c, \tau) = 1\{t - t_c = \tau\}$, with t_c being the first treatment year of cluster c (2010 or 2011) and τ the time to treatment (from -6 to 8). This is a dynamic treatment approach in which we assume that treatment effects are heterogeneous with time to treatment τ and treatment year t_c. See for instance Callaway and Sant'Anna (2021) for staggered treatment effects and references therein. An ATT for each τ and t_c is estimated assuming the conditional parallel trends assumption holds based on a "never-treated" group to account for the fact that there may have been an anticipation of the policy. However, alternatively assuming parallel trends based on a "not-yet-treated" group, that is assuming no treatment anticipation, essentially yields similar estimates. Figure 4 plots estimated values of μ_τ (aggregated over the two treatment years). It shows that there are no pretrends before treatment (when $\tau < 0$), supporting our identification strategy. We then observe a gradual increase from year to year. By the end of our observation period, the increase is on the order of 30% compared to the control group.

As explained in Section 3.2, we consider, in columns (2) to (4) of Table 3, several other specifications that address potential challenges to our main approach. Column (2) shows that the results are very similar when we remove individual fixed effects but add fixed effects leading to essentially comparing clusters with the same grade on research potential, the same scientific field and the same year of application. In column (3) we impose a restriction to rule out spillovers, and show that the effect is similar, although more imprecisely estimated, due to the smaller sample size. Finally in column (4), when we remove units that appear not linked to the central theme of the cluster, we find that the effect is also similar.

In the Supplementary Appendix, we consider two sets of robustness checks. First in Table A.4 we apply our main analysis to other outcomes: the number of publications involving

at least another coauthor from the same cluster, the number of publications involving only coauthors outside the cluster, creation of links that had not been formed previously and the intensity of within cluster collaboration. All the results are consistent with our main table: the researchers of funded clusters write more articles, have more collaborations and develop new links with other members of the cluster with whom they have never created in the past, and simultaneously write fewer papers with exclusively external collaborators.[25] In Table A.5 we use the same dependent variable as in our main table, but apply different specifications. We use Coarsened Exact Matching techniques, apply more stringent restrictions on the overall grade and finally restrict the sample based only on the grade specific to the goal of the project rather than the overall grade, and assume dynamic treatment effects. In all specifications, the results are close to our main specification.

4.2 Impact on Scientific Productivity

The previous subsection provides clear evidence of a large reorganization of the collaboration habits induced by the policy. Does this, in turn, have an effect on the productivity of these research clusters and on the type of research conducted?

We first examine whether the research cluster policy had an impact on quantitative measures of scientific production. We use the same identification strategies as in Table 3, but use the number of publications weighted by the journal impact factor as dependent variable. Table 4 shows that although the effect tends to be positive, it is non-significant. Expressing the results in terms of the 95% confidence relative to the mean of the dependent variable, and using specification (1), the effect on the number of publications weighted by IF is in the interval $[-7.1\%, 24\%]$ (relative to the mean). Figure 5 shows the dynamic treatment evolution (estimation of Equation 2).

We conduct the same type of robustness exercises as for our main analysis on links. In the Supplementary Appendix, Table A.6 we perform the same regressions on the other outcome

[25] The corresponding dynamic graphs are presented in Figure A.3.

variables: the number of publications, the number of cites, the number of top 10% and top 5% most cited papers. No significant effects emerge. In Table A.7, we reproduce Table A.5, but using the number of publications weighted by IF as dependent variable. This also yields very similar results (the estimated coefficient of interest is sometimes positive and sometimes negative but always very close to zero).[26]

While the policy had a large and significant impact on collaborations within the cluster, the impact on classical measures of production is more difficult to establish. We point out that, overall, the budget was relatively small: if the funding was equally shared among all the researchers that we identify as cluster members, this would represent less than 30K euros per researcher over an 8-year period, a much smaller amount than the amount of the typical individual grant.

4.3 Core and Periphery: Collaborations and Research Focus

Next, we explore the heterogeneity of the policy's impact between core and periphery members. In Table 5, we focus on the impact of the treatment on the number of links with core members (*LinksCore*) in column (1) and with periphery members (*LinksCore*) in column (2). Panel A presents the results for the core members and Panel B for periphery members. Columns (1) and (2) in both panels show that both core and periphery members increase collaborations on average, but that only the effect on links between core and periphery members is significant. In absolute terms, the effects are twice as large for core members, but in relative terms the links of periphery members with core members increase by 33% (links of the core with periphery increase by 19%).[27]

Overall this draws a picture of increased collaborations between core and periphery members, as we have observed in the two specific clusters described above (Figures 2 and 3). We see however, in the third column of Table 5, that those collaborations do not obviously

[26]The impact is positive and (barely) significant only in the last column of the table, where treatment is assumed dynamic, with an estimated 7.7% impact, relative to the mean dependent variable.

[27]Table A.8 in the Appendix reproduces the same regression results but excludes labs having no core member and yields similar results.

translate into more articles of periphery members. The coefficient is positive but not significant (in the band $[-6\%, 13\%]$). However, funding a research cluster organized around a theme may, rather than affecting global productivity, lead to a shift in the research focus of the cluster's members. The increase in collaborations between core and periphery members suggests a process whereby periphery members move the focus of their research toward the core theme of the research cluster.

To explore this potential impact, we use *ClusterTheme* as the dependent variable which is the number of papers using at least one keyword $\times$ subject category of the cluster universe, published in a given year. In the last column of Table 5, we see that the effects are not significantly different from zero for core members. We note, however, that for this group, the results are not easy to interpret, since some of their papers prior to the treatment are, by definition, used to define the set of cluster keywords and, thus, there is a natural tendency to underestimate the effect. The effects for members of the periphery do not suffer from this problem. We see in Panel B that this set of researchers appears to move closer to the cluster's research theme as they publish more papers using at least one keyword within the cluster universe. Specifically, they publish .084 more papers in the theme per year, which represents a 22% increase with respect to the average in this group.

5 Discussion and Interpretation

Our paper shows that even though the funding was relatively modest, the program had a large impact on the network of collaborations. The probability of having a within-cluster collaboration increased by 30% toward the end of the sample period, due to the funding. In this section we discuss mechanisms that explain these results and, in particular, justify why these newly formed collaborations, if productive, had not been formed previously.

To fix ideas, we conducted a small survey of 12 leading members of 10 funded clusters[28] and asked them to describe their perception of the effect of the funding. All of them stated

[28] 1 in math, 1 in social sciences, 3 in engineering, 2 in biology, 1 in genetics and 2 in physics.

that they witnessed either the creation (in 6 out of 10) or strengthening (in 4 out of 10) of links between group members. They also mentioned that common seminars were created involving several founding members of the research cluster, cosupervisions of PhD students were established and students typically shared common facilities.

The first mechanism that we argue can explain our results is one based on public good provision. The formation of the cluster decreased the cost of providing the public goods mentioned in the survey (seminars, cosupervision of PhDs, training), by, for instance, making it easier to organize seminars and training. More importantly, the cluster increased the benefits of providing these public goods by making contributions more visible and thus increasing the reputational benefits attached to them. These public goods, in turn, increased the value of internal collaborations. This mechanism can thus lead to the observed restructuring of the network of collaborations and can also explain why the periphery members, who are the beneficiaries of these public goods, have more to gain from the cluster policy, than the core members, who are providers.

These ideas are formalized in a theoretical framework presented in the Appendix. We model a potential cluster with two core members and two periphery members who optimally choose whether to work with members of the potential cluster or whether to build outside collaborations. Local public goods, which are simultaneously provided by the core members, increase the quality of potential internal collaborations. We model the creation of the cluster as decreasing the cost of providing the public good. Given this effect of the cluster, we show that in the Nash Equilibrium of the public good provision stage, the creation of the cluster increases the provision of public goods and as a consequence, increases the number of internal collaborations.

The second mechanism is based on the internal governance of these clusters. The newly formed entities had to create their own rules since no explicit guidelines were given by the funding agency regarding this dimension. Many chose to set up internal calls for funding where one of the requirements was to have members of at least two of the founding labs

involved in the project. This requirement would naturally lead to an increase in collaborations. Without knowing the objectives of the coordinators of these clusters, it is not easy to determine why these requirements for funding were explicitly introduced. Choosing a method to distribute funds to the members of the cluster was a difficult issue: it needed to guarantee a fair distribution of funds across groups or to at least minimize the perception of unfairness. At the same time it would have been too costly in time and effort to set up a full-fledged project evaluation program. One way of achieving a fair distribution while minimizing costs is to require that projects have at least two labs involved. This is also the approach chosen by many funding agencies, for instance, at the European level, where many funding instruments require the participation of several member countries. Both of these mechanisms, which are not mutually exclusive, can explain the increase in collaborations within the cluster.

6 Conclusion

Overall, our results suggest that the recent evolution, particularly in European countries, of moving toward policy instruments that fund research clusters rather than provide individual grants, tends to increase connections among local researchers and to shift research focus of cluster members. It may have been natural to expect that those who would benefit the most from the funding, would be the core members as their research lies at the forefront of the cluster's research theme. In fact, policy-makers may even have worried that the funds would be captured by a small group of core members. We show, on the contrary, that the effect of funding is larger for periphery members who collaborate more with core members and move their research direction toward the cluster's research theme in response to the policy. Given its low per capita cost, this policy based on local interactions seems to be efficient in inducing scientists to change their research direction whereas more traditional policy tools have recently been highlighted to have high switching costs (Myers, 2020). Whether these

organizational shifts will lead to productivity gains and contribute to excellence in the future remains an open question.

References

Aghion, Philippe and Peter Howitt. 1992. "A Model of Growth Through Creative Destruction." *Econometrica* 60(3):323–351.

Azoulay, Pierre and Danielle Li. 2020. "Scientific grant funding." *NBER Working Paper* (w26889).

Azoulay, Pierre, Joshua S Graff Zivin and Gustavo Manso. 2011. "Incentives and creativity: evidence from the academic life sciences." *The RAND Journal of Economics* 42(3):527–554.

Azoulay, Pierre, Joshua S. Graff Zivin and Jialan Wang. 2010. "Superstar extinction." *Quarterly Journal of Economics* 125(2):549–589.

Banal-Estañol, Albert, Inés Macho-Stadler and David Pérez-Castrillo. 2019. "Evaluation in research funding agencies: Are structurally diverse teams biased against?" *Research Policy* 48(7):1823–1840.

Borjas, George J. and Kirk B. Doran. 2012. "The collapse of the Soviet Union and the productivity of American mathematicians." *Quarterly Journal of Economics* 127(3):1143–1203.

Borjas, George J. and Kirk B. Doran. 2015. "Which peers matter? The relative impacts of collaborators, colleagues and competitors." *Review of Economics and Statistics* 97(5):1104–1117.

Bosquet, Clement, Pierre-Philippe Combes, Emeric Henry and Thierry Mayer. 2022. "Peer Effects in Academic Research: Senders and Receivers." *Economic Journal* 132:2644–2673.

Boudreau, Kevin J, Eva C Guinan, Karim R Lakhani and Christoph Riedl. 2016. "Looking across and looking beyond the knowledge frontier: Intellectual distance, novelty, and resource allocation in science." *Management Science* 62(10):2765–2783.

Bryan, Kevin and Heidi Williams. 2021. "Innovation: Market Failures and Public Policies." *Handbook of Industrial Organization* 5:281–388.

Callaway, Brantly and Pedro H. C. Sant'Anna. 2021. "Difference-in-Differences with multiple time periods." *Journal of Econometrics* 225(2):200–230.

Carayol, Nicolas and Marianne Lanoe. 2019. "The Design And The Impact of Project Funding In Science: Lessons From The ANR Experience." *Mimeo* .

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

Fang, Ferric C., Anthony Bowen and Arturo Casadevall. 2016. "Research: NIH peer review percentile scores are poorly predictive of grant productivity." *eLife Sciences.* 5:e13323.

Ginther, Donna K, Walter T Schaffer, Joshua Schnell, Beth Masimore, Faye Liu, Laurel L Haak and Raynard Kington. 2011. "Race, ethnicity, and NIH research awards." *Science* 333(6045):1015–1019.

Iacus, S. M., G. King and G. Porro. 2012. "Causal inference without balance checking: Coarsened exact matching." *Political analysis* 20(1):1–24.

Iaria, Alessandro, Carlo Schwarz and Fabian Waldinger. 2018. "Frontier Knowledge and Scientific Production: Evidence from the Collapse of International Science." *Quarterly Journal of Economics* 133(2):927–991.

Jacob, Brian A and Lars Lefgren. 2011. "The impact of research grant funding on scientific productivity." *Journal of public economics* 95(9-10):1168–1177.

Jaravel, Xavier, Neviana Petkova and Alex Bell. 2018. "Team-specific capital and innovation." *American Economic Review* 108(4):1034–1073.

Langfeldt, Liv, Mats Benner, Gunnar Sivertsen, Ernst H Kristiansen, Dag W Aksnes, Siri Brorstad Borlaug, Hanne Foss Hansen, Egil Kallerud and Antti Pelkonen. 2015. "Excellence and growth dynamics: A comparative study of the Matthew effect." *Science and Public Policy* 42(5):661–675.

Li, D. 2017. "Expertise versus Bias in Evaluation: Evidence from the NIH." *American Economic Journal: Applied Economics* 9(2):60–92.

Li, Danielle and Leila Agha. 2015. "Big names or big ideas: Do peer-review panels select the best science proposals?" *Science* 348(6233):434–438.

Möller, Torger, Marion Schmidt and Stefan Hornbostel. 2016. "Assessing the effects of the German Excellence Initiative with bibliometric methods." *Scientometrics* 109(3):2217–2239.

Myers, Kyle. 2020. "The Elasticity of Science." *American Economic Journal: Applied Economics* 12(4):103–34.

Oettl, Alexander. 2012. "Reconceptualizing stars: Scientist helpfulness and peer performance." *Management Science* 58(6):1122–1140.

Park, H., JJ. Lee and B-C. Kim. 2015. "Project selection in NIH: A natural experiment from ARRA." *Research Policy* 44(6):1145–1159.

Reijnhoudt, L., R. Costas, Ed Noyons, K. Borner and A. Scharnhorst. 2014. "'Seed + expand': a general methodology for detecting publication oeuvres of individual researchers." *Scientometrics* 101:1403–1417.

Romer, Paul. 1990. "Endogenous Technological Change." *Journal of Political Economy* 98(5):71–102.

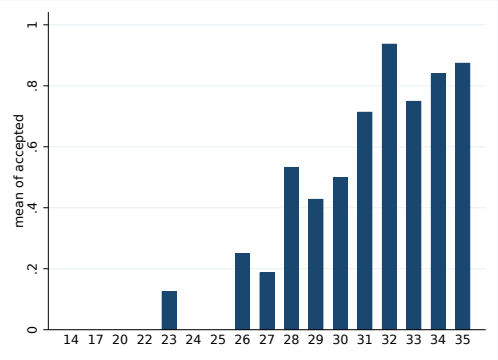
Waldinger, Fabian. 2012. "Peer effects in science: Evidence from the dismissal of scientists in Nazi Germany." *Review of Economic Studies* 79(2):838–861.

Wuchty, Stefan, Benjamin F Jones and Brian Uzzi. 2007. "The increasing dominance of teams in production of knowledge." *Science* 316(5827):1036–1039.

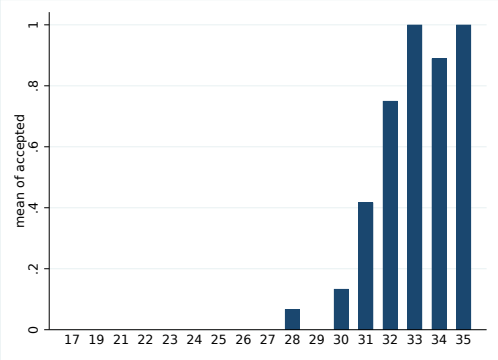
Figures and Tables

Figure 1: Probability of acceptance by grade

Panel A: LabEx selection as a function of overall grade

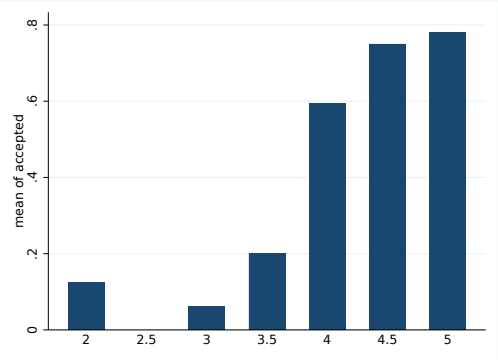


2010 contest

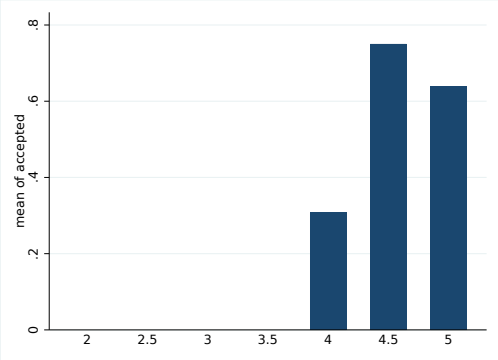


2011 contest

Panel B: LabEx selection as a function of grade on criterion 3 (goal of the project)



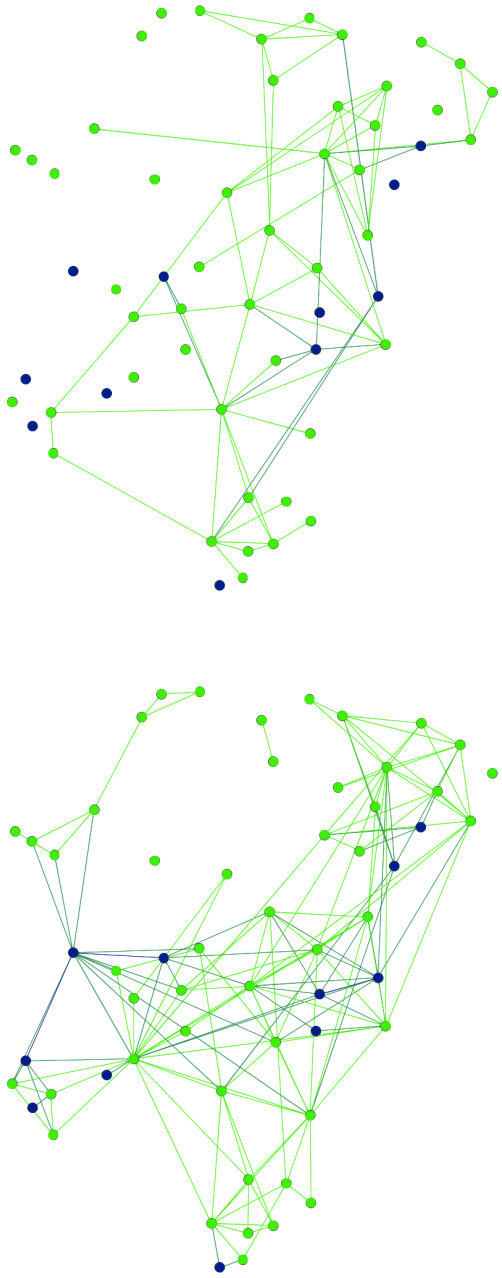
2010 contest



2011 contest

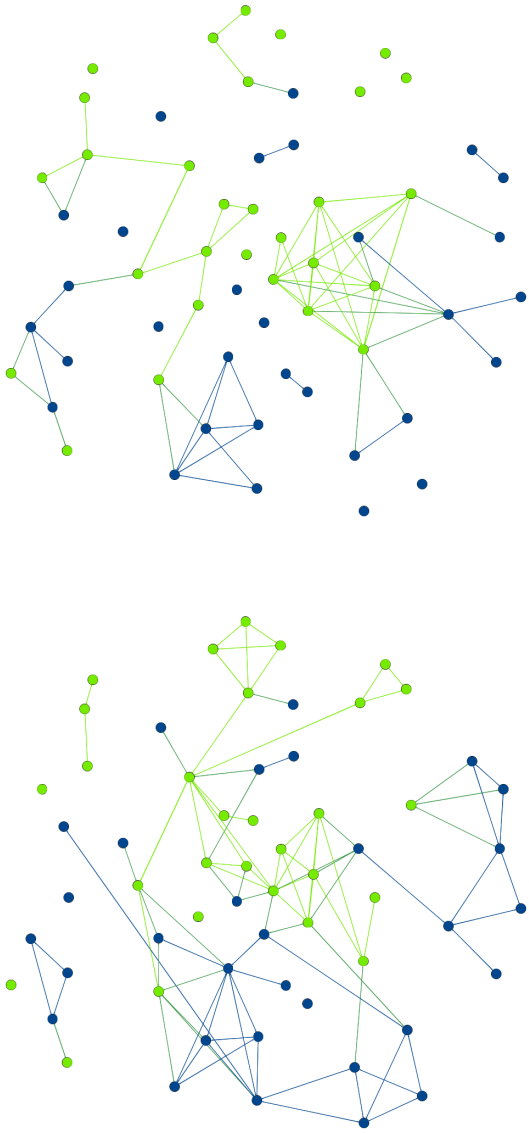
Notes: Panel A plots the proportion of clusters approved as a function of the total grade received by the research cluster (maximum grade is 35). Panel B plots the proportion of clusters approved as a function of the grade received on the 3rd criterion evaluating the goal of the project. In both panels the plots are produced separately for the 2010 and 2011 contests.

Figure 2: Network of collaborations for a treated cluster, before (top) and after treatment year (bottom)



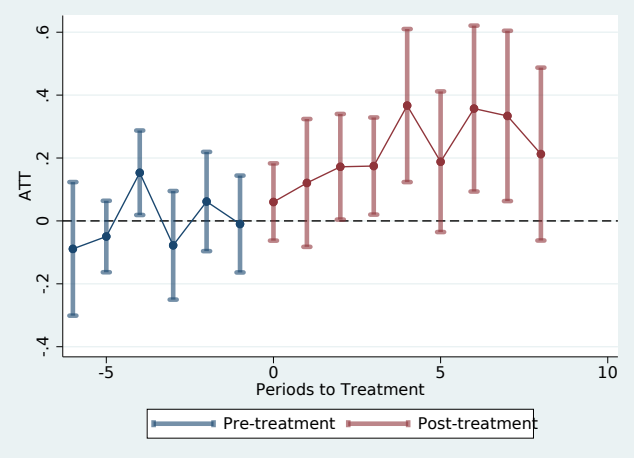
Notes: we represent the network of collaborations for a particular cluster. Before treatment graph on the top (all years <2012); after treatment graph at the bottom (all years ≥ 2012). Nodes represent researchers (we restrict to those who were part of a single proposal). Light green nodes represent core members and dark blue nodes are periphery members. An edge stands for at least one joint paper in each considered period. The color of each edge is a mixture of nodes' color at both ends.

Figure 3: Collaborations in a non treated research cluster proposals, before (top) and after treatment year (bottom)



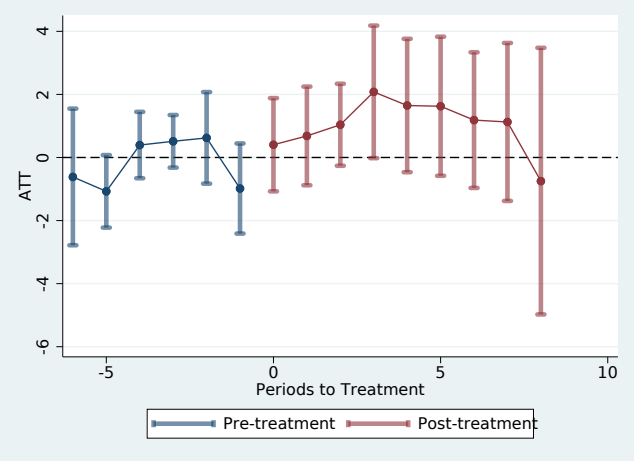
Notes: Same as Figure 2 but for a different cluster that was not funded.

Figure 4: Organization of research: comparing financed and non-financed clusters



Note: we estimate Equation 2, using Links as the dependent variable and applying to the data the restriction on overall grades (the sample is restricted to researchers who were part of a single cluster proposal that received an overall grade between 26 and 32 for the 2010 contest and between 30 and 32 for the 2011 contest). Following Callaway and Sant'Anna (2021) the average treatment effects are estimated separately for each cohort and aggregated to produce values of μ_τ that we plot. The upper and lower values of each bar are 95 percent confidence intervals of robust standard errors clustered at the research cluster level.

Figure 5: Production of research: comparing financed and non-financed clusters



Note: we estimate Equation 2, using AIF (number of articles weighted by the journal impact factor) as the dependent variable and applying to the data the restriction on overall grades (the sample is restricted to researchers who were part of a single cluster proposal that received an overall grade between 26 and 32 for the 2010 contest and between 30 and 32 for the 2011 contest). Following Callaway and Sant'Anna (2021) the average treatment effects are estimated separately for each cohort and aggregated to produce values of μ_τ that we plot. The upper and lower values of each bar are 95 percent confidence intervals of robust standard errors clustered at the research cluster level.

Table 1: Descriptive statistics on research clusters

	(1) Treated		(2) Non Treated		(3) Difference t-test	
	mean	sd	mean	sd	b	p
Number of scientists	304.15	289.23	253.41	272.26	-0.17	(0.08)
Number of research units	16.13	12.36	14.22	12.70	-0.12	(0.14)
Share of scientists in the core	0.31	0.24	0.32	0.24	0.03	(0.71)
Number of articles ($PubsRC$)	469.32	453.94	385.87	414.99	-0.18	(0.07)
Number of articles weighted by cites ($CitesRC$)	3204.43	3164.71	2655.03	3126.45	-0.17	(0.09)
Nbr of articles w. by Impact Factor ($AIFRC$)	2358.86	2317.79	1948.67	2251.34	-0.17	(0.09)
Field: biology	0.29	0.45	0.32	0.47	0.14	(0.46)
Field: chemistry	0.06	0.23	0.10	0.30	0.83	(0.09)
Field: physics	0.19	0.39	0.08	0.28	-0.58	(0.00)
Field: engineering	0.12	0.33	0.14	0.35	0.17	(0.55)
Field: mathematics	0.10	0.30	0.05	0.22	-0.5	(0.09)
Field: SSH	0.25	0.43	0.27	0.44	0.08	(0.61)
Overall grade	31.97	2.19	27.07	3.50	-0.15	(0.00)
Grade: team quality	4.75	0.65	4.09	0.85	-0.14	(0.00)
Grade: goal	4.52	0.74	3.72	0.91	-0.17	(0.00)
Grade: potential research	4.43	0.74	3.63	0.91	-0.18	(0.00)
Grade: potential training	4.54	0.75	3.95	0.85	-0.13	(0.00)
Grade: organization	4.33	0.78	3.72	0.84	-0.14	(0.00)
Grade: structure	4.58	0.72	3.94	0.88	-0.14	(0.00)
Grade: resource generation	4.42	0.77	3.65	0.90	-0.17	(0.00)
Observations	163		216		379	

Note: Column (1) presents the mean and sd of our main variables when we restricts the sample to research clusters that were financed and column (2) restricts to those that did were not funded. Column (3) presents the difference (standardized by the mean of column 1) between the means of columns (1) and (2) and the p value of a test of differences of means.

Table 2: Descriptive statistics on individual researchers: Full sample

	(1) Treated		(2) Non Treated		(3) Difference t-test	
	mean	sd	mean	sd	b	p
Age (*Age*)	41.24	10.45	42.07	10.30	0.02	(0.00)
Number of articles (*Pubs*)	2.28	3.41	1.98	3.35	-0.13	(0.00)
Adjusted number of articles weighted by cites (*Cites*)	53.37	150.72	45.66	133.91	-0.14	(0.00)
Number of articles weighted by Impact Factor (*AIF*)	12.72	23.93	10.76	20.84	-0.15	(0.00)
Number of links within the cluster (*Links*)	1.05	1.76	0.83	1.39	-0.21	(0.00)
Number of collaborative articles within the cluster (*CollaPubs*)	1.00	1.90	0.81	1.55	-0.19	(0.00)
Number of collaborations within the cluster (*Collaborations*)	2.07	5.32	1.42	3.65	-0.31	(0.00)
Number of external articles (*ExternalPubs*)	1.28	2.52	1.17	2.69	-0.07	(0.00)
Number of new links within the cluster (*NewLinks*)	0.37	0.63	0.28	0.48	-0.24	(0.00)
Observations	30409		10684		41093	

Note: All columns present the mean and sd of our main variables for the overall sample of researchers observed before 2010 to compare pre-treatment samples. Column (1) restricts the sample to researchers in research clusters that were financed and Column (2) restricts to those that were in research clusters proposals that did not receive financing. Column (3) presents the difference (standardized by the mean of column 1) between the means of columns (1) and (2) and the p value of a test of differences of means.

Table 3: Effects of cluster policy on within cluster collaborations

Model	(1)	(2)	(3)	(4)
Dependent variable	Links within the cluster (*Links*)			
Treatment × Post	0.221**	0.284**	0.219*	0.251***
	(0.086)	(0.117)	(0.109)	(0.082)
Selection on total grade	Yes	Yes	Yes	Yes
Individual fixed effects	Yes	No	Yes	Yes
Grade×Field×Treatment year FE	No	Yes	No	No
Excluding ≤10km radius & same field	No	No	Yes	No
Excluding labs with no core member	No	No	No	Yes
Observations	95259	95259	43856	86878
Number of Clusters	79	79	35	83
Mean dep variable	1.4	1.4	1.4	1.3
Adj. R-Square	.52	.13	.54	.51

*Note: In all columns the dependent variable measures the number of links within the cluster. In column (1) we present the results of the estimation of Equation 1, applying to the data the restriction on grades (the sample is restricted to researchers who were part of a single cluster proposal that received an overall grade between 26 and 32 for the 2010 contest and between 30 and 32 for the 2011 contest). Column (2) estimates a model replacing individual fixed effects with grade on criterion 3 × discipline × year of application fixed effects. Column (3) estimates the same model as in column (1) but restricting the data to clusters for which there is no other research cluster in their field within a 10 km radius. Column (4) estimates the same model as in column (1) but excluding from the data observations attached to research units where none of their researchers appear in the bibliography of the LabEx. Significance levels are given by: * p<0.1, ** p<0.05, *** p<0.01.*

Table 4: Effects of cluster policy on scientific outcomes

Model	(1)	(2)	(3)	(4)
Dependent variable	Articles weighted by IF (AIF)			
Treatment × Post	0.614	1.166	0.581	0.699
	(0.535)	(1.104)	(0.816)	(0.564)
Selection on total grade	Yes	Yes	Yes	Yes
Individual fixed effects	Yes	No	Yes	Yes
Grade×Field×Treatment year FE	No	Yes	No	No
Excluding ≤10km radius & same field	No	No	Yes	No
Excluding labs with no core member	No	No	No	Yes
Observations	95259	95259	43856	86758
Number of Clusters	79	79	35	83
Mean dep variable	14	14	13	14
Adj. R-Square	.59	.11	.64	.59

*Note: In all columns the dependent variable measures the number of publications weighted by the journal impact factor. In column (1) we present the results of the estimation of Equation 1, applying to the data the restriction on grades (the sample is restricted to researchers who were part of a single cluster proposal that received an overall grade between 26 and 32 for the 2010 contest and between 30 and 32 for the 2011 contest). Column (2) estimates a model replacing individual fixed effects with grade on criterion 3 × discipline × year of application fixed effects. Column (3) estimates the same model as in column (1) but restricting the data to clusters for which there is no other research cluster in their field within a 10 km radius. Column (4) estimates the same model as in column (1) but excluding from the data observations attached to research units where none of their researchers appear in the bibliography of the LabEx. Significance levels are given by: * p<0.1, ** p<0.05, *** p<0.01.*

Table 5: Cluster policy on core vs. periphery members

Panel A: Core members

Model	(1)	(2)	(3)	(4)
Dependent variable	*LinksCore*	*LinksPeriph*	*AIF*	*ClusterTheme*
Treatment × Post	0.143	0.158**	1.667	0.006
	(0.119)	(0.069)	(1.453)	(0.084)
Observations	24418	24418	24418	22507
Number of Clusters	69	69	69	56
Mean dep variable	1.1	.82	20	1.2
Adj. R-Square	.51	.5	.59	.55

Panel B: Periphery members

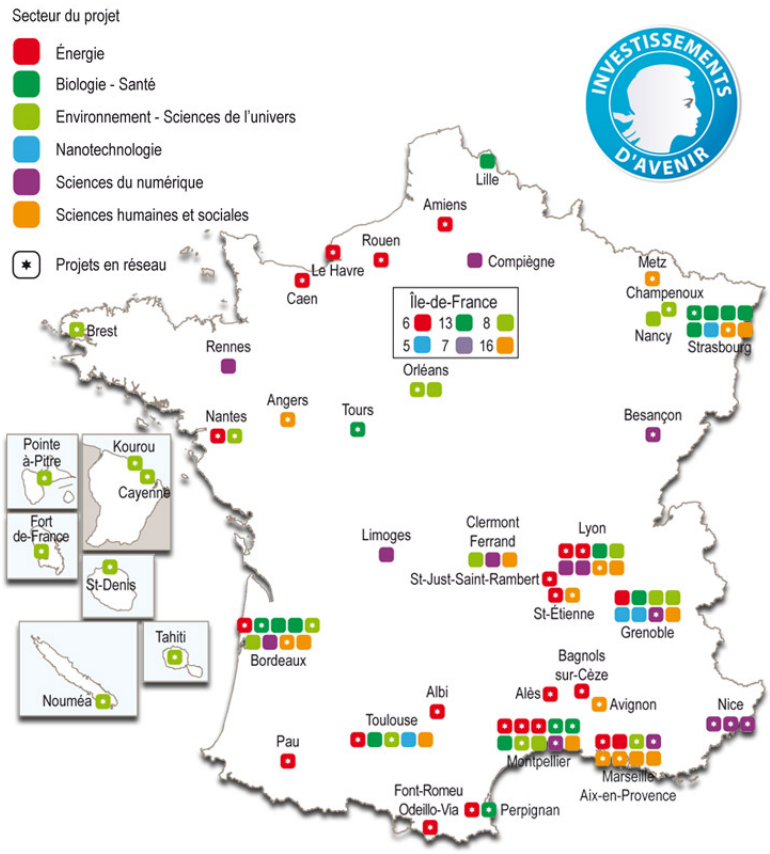
Model	(1)	(2)	(3)	(4)
Dependent variable	*LinksCore*	*LinksPeriph*	*AIF*	*ClusterTheme*
Treatment × Post	0.075***	0.129	0.415	0.084***
	(0.024)	(0.089)	(0.576)	(0.029)
Observations	70771	70771	70771	53027
Number of Clusters	78	78	78	57
Mean dep variable	.23	.92	12	.39
Adj. R-Square	.43	.5	.59	.52

*Note: In Panel A, the sample is restricted to members of the core, while in Panel B it is restricted to members of the periphery. In column (1) the dependent variable is the number of links with members of the core, whereas in column (2) the dependent variable is the number of links with members of the periphery. In column (3) the dependent variable is the number of articles weighted by the journal impact factor. In column (4) the dependent variable is the number of papers using at least one keyword × subject category in the cluster universe, published in a given year. In both panels, we report coefficients estimated on Equation 1 applying to the data the restriction on grades (the sample is restricted to researchers who were part of a single cluster proposal that received an overall grade between 26 and 32 for the 2010 contest and between 30 and 32 for the 2011 contest). Standard errors are clustered at the research cluster level. Significance levels are given by: * p<0.1, ** p<0.05, *** p<0.01.*

Appendix (For Online Publication)

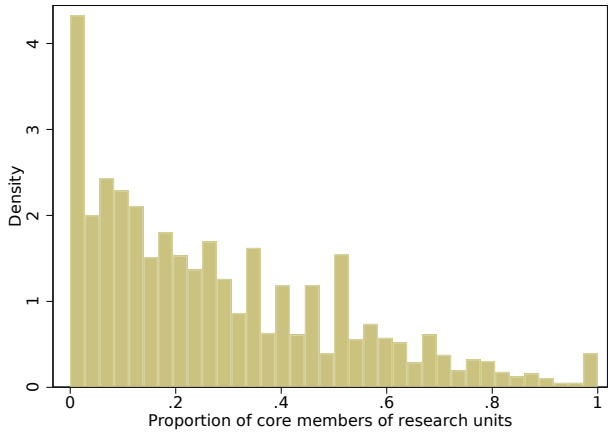
Geography of the program

Figure A.1: Map of awarded research clusters by field of study



This map was obtained from https : //media.enseignementsup−recherche.gouv.fr/file/Fiches$_L$*abex/05/0/IA*$_L$*ABEX*$_C$*arte*$_V$*def*$_1$*72050.pdf.*

Figure A.2: Distribution of the share of individuals in the core per unit-cluster pair



Notes: for each unit-cluster pair we compute the share of researchers in the unit who are in the bibliography of the cluster. The figure plots the distribution of these shares.

Table A.1: Descriptive statistics on individual researchers: Restriction on overall grade

	(1) Treated		(2) Non Treated		(3) Difference t-test	
	mean	sd	mean	sd	b	p
Age (*Age*)	41.47	10.53	41.31	9.91	-0.01	(0.59)
Number of articles (*Pubs*)	2.09	3.22	2.31	3.37	0.11	(0.02)
Adjusted number of articles weighted by cites (*Cites*)	50.94	144.99	53.82	156.64	0.06	(0.50)
Number of articles weighted by Impact Factor (*AIF*)	11.94	22.52	12.20	23.35	0.02	(0.68)
Number of links within the cluster (*Links*)	0.94	1.56	1.07	1.57	0.14	(0.00)
Number of collaborative articles within the cluster (*CollaPubs*)	0.88	1.66	1.10	1.89	0.25	(0.00)
Number of collaborations within the cluster (*Collaborations*)	1.72	4.49	1.85	3.40	0.07	(0.23)
Number of external articles (*ExternalPubs*)	1.21	2.51	1.21	2.38	0.00	(0.95)
Number of new links within the cluster (*NewLinks*)	0.34	0.54	0.36	0.53	0.03	(0.43)
Observations	5368		1675		7043	

Note: The sample is limited to researchers who were part of a single cluster proposal that received an overall grade between 26 and 32 for the 2010 contest and between 30 and 32 for the 2011 contest. All columns present the mean and sd of our main variables observed before 2010 to compare pre-treatment samples. Column (1) restricts the sample to researchers in research clusters that were financed and Column (2) restricts to those that were in research clusters proposals that did not receive financing. Column (3) presents the difference (standardized by the mean of column 1) between the means of columns (1) and (2) and the p value of a test of differences of means.

Table A.2: Descriptive statistics on individual researchers: core vs. periphery members

Panel A: Full sample

	(1) Core		(2) Periphery		(3) Difference t-test	
	mean	sd	mean	sd	b	p
Age (*Age*)	42.77	9.99	41.03	10.52	-0.04	(0.00)
Number of articles (*Pubs*)	3.02	3.82	1.93	3.21	-0.36	(0.00)
Number of articles weighted by citations (*Cites*)	78.40	182.78	42.49	131.34	-0.45	(0.00)
Number of articles weighted by Impact Factor (*AIF*)	17.94	26.40	10.33	21.69	-0.42	(0.00)
Number of links within the cluster (*Links*)	1.59	2.14	0.79	1.44	-0.50	(0.00)
Number of collaborative articles within the cluster (*CollaPubs*)	1.52	2.26	0.76	1.60	-0.50	(0.00)
Number of collaborations within the cluster (*Collaborations*)	3.00	5.62	1.54	4.66	-0.49	(0.00)
Number of external articles (*ExternalPubs*)	1.50	2.65	1.17	2.53	-0.22	(0.00)
Number of new links within the cluster (*NewLinks*)	0.54	0.72	0.29	0.54	-0.46	(0.00)
Observations	10158		30935		41093	

Panel B: Restriction on overall grade

	(1) Core		(2) Periphery		(3) Difference t-test	
	mean	sd	mean	sd	b	p
Age (*Age*)	42.75	9.83	41.31	9.91	-0.04	(0.00)
Number of articles (*Pubs*)	3.01	3.73	2.31	3.37	-0.38	(0.00)
Number of articles weighted by citations (*Cites*)	88.22	210.21	53.82	156.64	-0.55	(0.00)
Number of articles weighted by Impact Factor (*AIF*)	18.21	26.37	12.20	23.35	-0.45	(0.00)
Number of links within the cluster (*Links*)	1.59	2.01	1.07	1.57	-0.52	(0.00)
Number of collaborative articles within the cluster (*CollaPubs*)	1.54	2.24	1.10	1.89	-0.52	(0.00)
Number of collaborations within the cluster (*Collaborations*)	2.84	5.01	1.85	3.40	-0.51	(0.00)
Number of external articles (*ExternalPubs*)	1.47	2.50	1.21	2.38	-0.23	(0.00)
Number of new links within the cluster (*NewLinks*)	0.55	0.67	0.36	0.53	-0.49	(0.00)
Observations	1769		1675		7043	

Note: Panel A includes the full sample, while Panel B applies the restriction on grades (the sample is restricted to researchers who were part of a single cluster proposal that received an overall grade between 26 and 32 for the 2010 contest and between 30 and 32 for the 2011 contest). We restrict our data to values before 2010 to compare the pre-treatment samples. Column (1) presents the mean and sd of our main variables for the sample of core researchers and Column (2) restricts to Periphery members. Column (3) presents the difference (standardized by the mean of column 1) between the means of columns (1) and (2) and the p value of a test of differences of means.

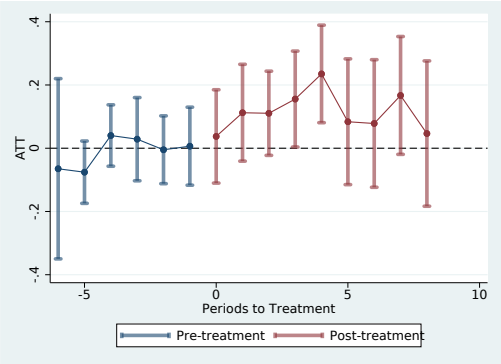
Table A.3: Comparing researchers in one vs. in two or three clusters

	(1)		(2)		(3)	
	1 cluster		2 or 3 clusters		Difference t-test	
	mean	sd	mean	sd	b	p
Number of articles (*Pubs*)	2.07	3.38	2.11	3.25	-0.02	(0.27)
Number of articles weighted by cites (*Cites*)	49.60	147.05	46.74	136.17	0.05	(0.07)
Number of articles weighted by the journal impact factor (*AIF*)	11.77	22.76	11.21	21.99	0.05	(0.02)
Mean number of authors (*TeamSize*)	9.98	28.70	9.44	27.25	0.05	(0.12)
Number of collaborative articles within the cluster (*CollaPubs*)	0.86	1.80	0.91	1.77	-0.05	(0.04)
Number of links within the cluster (*Links*)	0.91	1.66	0.93	1.63	-0.02	(0.29)
Number of collaborations within the cluster (*Collaborations*)	1.80	5.57	1.79	4.76	0.01	(0.87)
Number of external articles (*ExternalPubs*)	1.21	2.60	1.21	2.40	0.00	(0.97)
Number of new links within the cluster (*NewLinks*)	0.32	0.57	0.32	0.56	-0.00	(0.83)
Observations	14982		18510		33492	

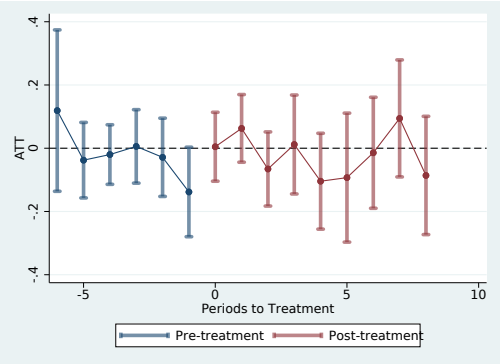
Note: we restrict our data to values before 2010 to compare the pre-treatment samples. Column (1) presents the mean and sd of our main variables for researchers who are part of only one cluster proposal and Column (2) restricts to those who are part of two or three cluster proposals. Column (3) presents the difference (standardized by the mean of column 1) between the means of columns (1) and (2) an the p value of a test of differences of means.

Robustness

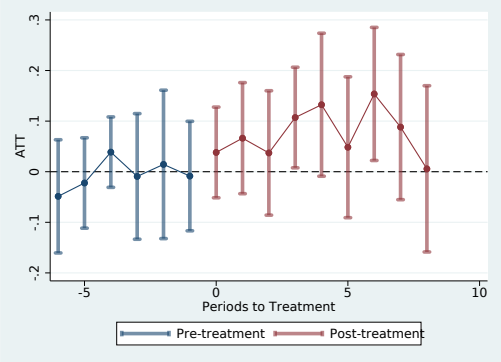
Figure A.3: Organization of research: additional outcomes



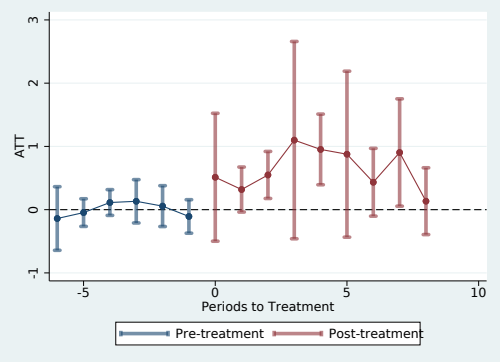
(a) Number of publications with a co-author in cluster



(b) Number of publications with no co-author in cluster



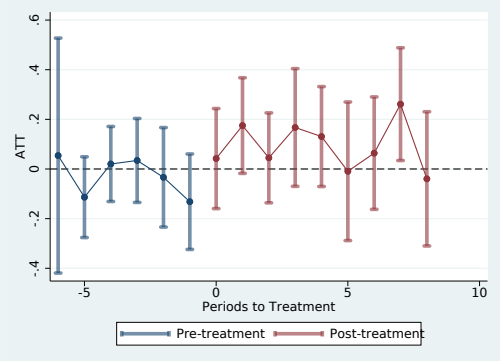
(c) Number of new links same cluster



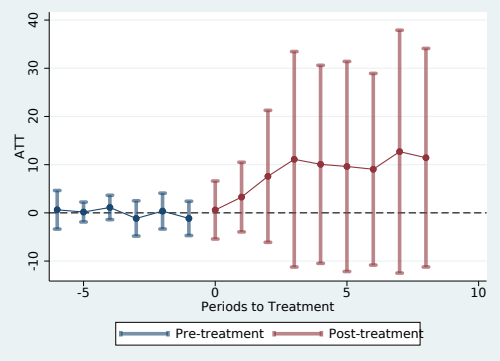
(d) Number of collaborations same cluster

Note: we estimate specification in Equation 2, using different dependent variables nd applying to the data the restriction on overall grades (the sample is restricted to researchers who were part of a single cluster proposal that received an overall grade between 26 and 32 for the 2010 contest and between 30 and 32 for the 2011 contest). Following (Callaway and Sant'Anna, 2021) the average treatment effects are estimated separately for each cohort and aggregated to produce values of μ_τ that we plot. The upper and lower values of each bar are 95 percent confidence intervals of robust standard errors clustered at the research cluster level. In Panel (a), the dependent variable measures the number of publications with at least one coauthor from the cluster. In Panel (b), the dependent variable measures the number of publications with at not a single coauthor from the cluster. In Panel (c), the dependent variable measures the number of new links with other members of the cluster that had never been formed in the past. In Panel (d), the dependent variable measures the number of collaborations with a coauthor from the cluster.

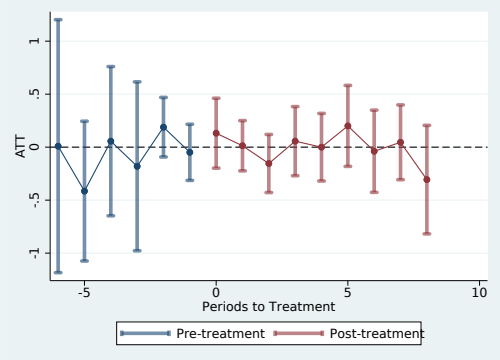
Figure A.4: Production of research: additional outcomes



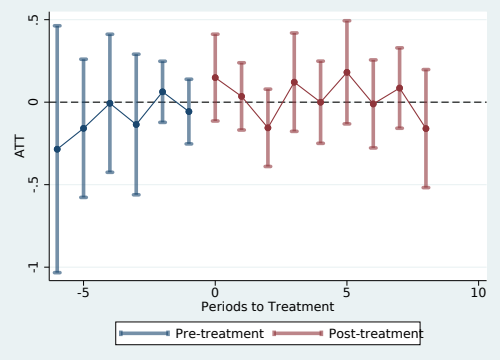
(a) Number of publications



(b) Number of cites



(c) Number of publications in the top 10%



(c) Number of publications in the top 5%

Note: we estimate specification in Equation 2, using different dependent variables and applying to the data the restriction on overall grades (the sample is restricted to researchers who were part of a single cluster proposal that received an overall grade between 26 and 32 for the 2010 contest and between 30 and 32 for the 2011 contest). Following Callaway and Sant'Anna (2021) the average treatment effects are estimated separately for each cohort and aggregated to produce values of μ_τ that we plot. The upper and lower values of each bar are 95 percent confidence intervals of robust standard errors clustered at the research cluster level. In Panel (a), the dependent variable measures the number of articles published. In Panel (b), the papers are weighted by cites in a 3-year window. In Panel (c), only papers that are in the top 10% most cited in their subject category are counted. In Panel (d), only those in the top 5% are considered.

Table A.4: Effects of the cluster policy on other outcomes measuring organization

Model	(1)	(2)	(3)	(4)
Dependent variable	*CollaPubs*	*ExternalPubs*	*NewLinks*	*Collaborations*
Treatment × Post	0.147**	-0.101*	0.082***	0.595**
	(0.057)	(0.057)	(0.029)	(0.256)
Observations	95259	95259	95259	95259
Number of Clusters	79	79	79	79
Mean dep variable	1.2	1.3	.43	2.5
Adj. R-Square	.55	.61	.17	.4

*Note: In all columns the results are obtained using Equation 1, applying to the data the restriction on grades (the sample is restricted to researchers who were part of a single cluster proposal that received an overall grade between 26 and 32 for the 2010 contest and between 30 and 32 for the 2011 contest). In column (1), the dependent variable measures the number of publications with at least one coauthor from the cluster. In column (2), the dependent variable measures the number of publications with no coauthor from the cluster. In column (3), the dependent variable measures the number of new links with a member of a cluster, i.e. links that had never been formed in the past. In column (4), the dependent variable measures the number of collaborations with a coauthor from the cluster (each co-author on a given paper is a collaboration here). We include individual and time fixed effects. The standard errors are clustered at the research cluster level. Significance level are given by: * p<0.1, ** p<0.05, *** p<0.01.*

Table A.5: Effects of the cluster policy on within cluster collaborations : alternative restrictions

Model	(1)	(2)	(3)	(4)
Dependent variable	Links within the cluster (*Links*)			
Treatment × Post	0.182**	0.310**	0.184**	0.220***
	(0.072)	(0.133)	(0.076)	(0.079)
Selection on total grade	Yes	No	No	Yes
Coarsened exact matching	Yes	No	No	No
Alternative selection on total grade	No	Yes	No	No
Selection on criterion 3 grade	No	No	Yes	No
Dynamic treatment assumption	No	No	No	Yes
Observations	49198	44711	85074	95325
Number of Clusters	67	30	90	79
Mean dep variable	1	1.4	1.3	1.2
Adj. R-Square	.47	.51	.5	

*Note: we reproduce Table 3 but changing the restrictions or adding weights on observations. In column (1), we apply the same restrictions as in the main regression (a unique cluster participation and restrictions on grades) and we also apply weights calculated from a Coarsened Exact Matching performed using the following variables: exact total grade received and field of specialization of the cluster, as well as some individual variables such as age, the number of articles and the number of cites in the last three years. In column (2), we change the restriction on grades. Specifically we restrict the approval probability to be between 40% and 60% which implies keeping grades between 28 and 31 in 2010 and equal to 31 in 2011. In column (3), we do not impose the same restriction on grades as in Table 3 but we restrict the sample to researchers who participated in a single cluster application that received a grade of 4 for the goals of the project (criterion 3). In column (4), the ATT is first estimated as for our main estimation, but separateley for each year and each time to treatment, and then aggregated. In all the panels, we include individual and time fixed effects and cluster the standard errors at the research cluster level. Significance levels are given by: * p<0.1, ** p<0.05, *** p<0.01.*

Table A.6: Effects of the cluster policy on other outcomes measuring productivity

Model	(1)	(2)	(3)	(4)
Dependent variable	*Pubs*	*Cites*	*Top10*	*Top5*
Treatment × Post	0.046	10.077	-0.158	-0.071
	(0.072)	(8.822)	(0.195)	(0.105)
Observations	95259	95259	86959	86959
Number of Clusters	79	79	79	79
Mean dep variable	2.5	89	2	1.1
Adj. R-Square	.64	.75	.37	.31

*Note: we present the results of the estimation of Equation 1, for outcome variables measuring connections, applying to the data the restriction on grades (the sample is restricted to researchers who were part of a single cluster proposal that received an overall grade between 26 and 32 for the 2010 contest and between 30 and 32 for the 2011 contest). In column (1), the dependent variable measures the number of publications. In column (2), articles are weighted by citations made in the 3-year period. In column (3), only articles that are among the top 10% most cited in their subject category are counted. In column (4), only those in the top 5% are considered. We include individual and time fixed effects. The standard errors are clustered at the research cluster level. Significance levels are given by: * p<0.1, ** p<0.05, *** p<0.01.*

Table A.7: Effects of cluster policy on productivity: alternative restrictions

Model	(1)	(2)	(3)	(4)
Dependent variable	Articles weighted by IF (AIF)			
Treatment × Post	-0.305	-0.003	-0.146	1.08*
	(0.491)	(0.966)	(0.511)	(0.64)
Selection on total grade	Yes	No	No	Yes
Coarsened exact matching	Yes	No	No	No
Alternative selection on total grade	No	Yes	No	No
Selection on criterion 3 grade	No	No	Yes	No
Dynamic treatment assumption	No	No	No	Yes
Observations	49198	44711	85074	95325
Number of Clusters	67	30	90	79
Mean dep variable	10	14	14	14
Adj. R-Square	.51	.55	.59	

*Note: we reproduce Table 4 but changing the restrictions or adding weights on observations. In column (1), we apply the same restrictions as in the main regression (a unique cluster participation and restrictions on grades) and we also apply weights calculated from a Coarsened Exact Matching performed using the following variables: exact total grade received and field of specialization of the cluster, as well as some individual variables such as age, the number of articles and the number of cites in the last three years. In column (2), we change the restriction on grades. Specifically we restrict the approval probability to be between 40% and 60% which implies keeping grades between 28 and 31 in 2010 and equal to 31 in 2011. In column (3), we do not impose the same restriction on grades as in Table 4 but we restrict the sample to researchers who participated in a single cluster application that received a grade of 4 for the goals of the project (criterion 3). In column (4), the ATT is first estimated as for our main estimation, but separateley for each year and each time to treatment, and then aggregated. In all the panels, we include individual and time fixed effects and cluster the standard errors at the research cluster level. Significance levels are given by: * p<0.1, ** p<0.05, *** p<0.01.*

Table A.8: Effects of the cluster policy on core vs. periphery: restricting to active units

Panel A: Core members

Model	(1)	(2)	(3)	(4)
Dependent variable	*LinksCore*	*LinksPeriph*	*AIF*	*ClusterTheme*
Treatment × Post	0.124	0.156**	1.279	-0.003
	(0.111)	(0.063)	(1.360)	(0.086)
Observations	26155	26155	26155	23239
Number of Clusters	80	80	80	56
Mean dep variable	1.1	.8	20	1.2
Adj. R-Square	.51	.5	.59	.54

Panel B: Periphery members

Model	(1)	(2)	(3)	(4)
Dependent variable	*LinksCore*	*LinksPeriph*	*AIF*	*ClusterTheme*
Treatment × Post	0.088***	0.175**	0.522	0.100***
	(0.028)	(0.077)	(0.606)	(0.028)
Observations	60534	60534	60534	50367
Number of Clusters	81	81	81	56
Mean dep variable	.28	.82	12	.39
Adj. R-Square	.41	.47	.58	.52

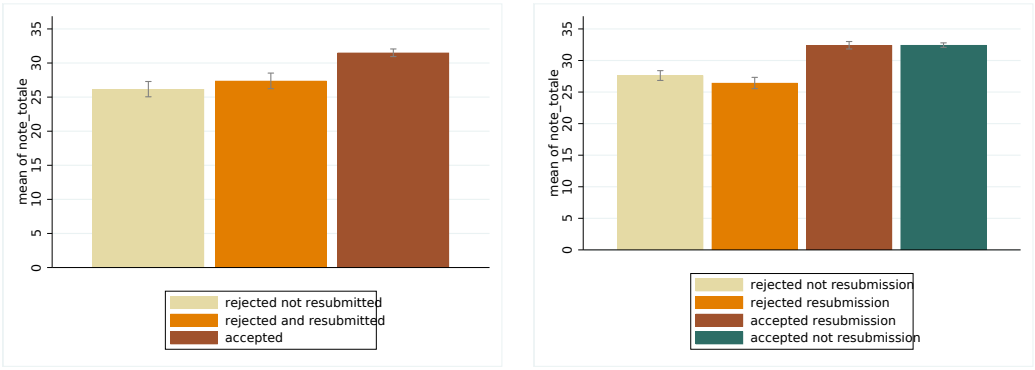
*Note: In Panel A, the sample is restricted to members of the core, while in Panel B it is restricted to members of the periphery. In column (1) the dependent variable is the number of links with members of the core, whereas in column (2) the dependent variable is the number of links with members of the periphery. In column (3) the dependent variable is the number of articles weighted by the journal impact factor. In column (4) the dependent variable is the number of papers using at least one keyword × subject category in the cluster universe, published in a given year. In both panels, we report coefficients estimated on Equation 1 applying to the data the restriction on grades (the sample is restricted to researchers who were part of a single cluster proposal that received an overall grade between 26 and 32 for the 2010 contest and between 30 and 32 for the 2011 contest) and excluding from the data observations attached to research units when none of their researchers appear in the bibliography of the cluster proposal. Standard errors are clustered at the research cluster level. Significance levels are given by: * p<0.1, ** p<0.05, *** p<0.01.*

Projects resubmitted

As described in the main text, projects rejected in the 2010 edition, could be resubmitted in the 2011 one. Out of the 195 submitted projects in 2011, 55 were resubmissions, out of which 20 were funded. The resubmissions were typically significantly changed, with a change both in the number of partners and in the budget.

In Figure A.5 below, we compare the grades obtained by the resubmitted projects, with the others, both for the 2010 and 2011 editions. We see that, in the 2010 edition, the cluster proposals that were rejected and not resubmitted are not significantly different from those that were simply rejected and not resubmitted. Similarly, for the 2011 edition, both for rejected and accepted projects, there is no difference between the proposals that were re-submissions and those that were not.

Figure A.5: Average grade by type of LabEx



2010 contest 2011 contest

Notes: This figure presents the average of the total grade obtained by projects, separately for the 2010 and 2011 editions and distinguishing projects that were resubmitted versus those that were not.

Data Construction: Professors and Researchers

Data collection is eased in France thanks to the centralization of information at the national level through various procedures. Through bilateral contracting with the government and research institutes research units report every four-to-five years the list of their tenured research staff to the MHER. The ministry also maintains a list of all professors (Professeur des Universités) and assistant professors (Maîtres de Conférence) since 2000 as, though they are in practice employed by universities, those persons are formally civil servants and thus paid by the government. National research institutes like the CNRS, also maintain the information of all their employees and provided us that information. We had access to those sensitive data provided that the team abides to strict rules protecting personal data that are pseudo-anonymized and data security. As information on specific personal profiles come from observations at different points in time and sometimes from different sources, a huge manual and automatic disambiguation work has been performed. The disambiguation of individual profiles is easy when they scientists are associated to the same location, research unit and institution across tables. As research units evolve over time (either birth, extinct, merge or split), we used a very convenient national roster of research units in which research units have specific identifiers that are consolidated over time by all research institutes, the ministry and universities.

Researchers and professors may however move over their career or be simultaneously associated to different units. Then, partial information, such a birth dates, or even web search procedures were designed to obtain the best possible list of professors and researchers. Yearly observations are used to imply entry and exit dates of each person in each unit and from the sample. At the end of this task, we end out with a consistent roster of nearly 97,122 tenured researchers and professors affiliated to 256 universities and research institutes and 3,250 distinct research units.

We estimate this dataset gathers nearly 96% of the reference population at the national level over the considered period. Indeed, The MHER documents there are about 56,000

professors and assistant professors in France in 2015 and 17,000 tenured researchers in the national research institutes. From yearly counts of final exits (retirements and death) in those data, we estimate that about 28,000 more distinct persons (23,000 professors and 5,000 researchers) have been part of this population over the period 2005-2019. This leads to a raw estimate of the total population of professors and researchers over the period of about 101,000 persons. A few research units may not certified by the Ministry of Higher Education and Research (MHER). This happens when the research units are funded only by other ministries (Industry, Agriculture, Defense) via specific higher education schools or research institutes.

Theoretical framework

Model

Consider two core members $c \in \{1, 2\}$ and two periphery members $p \in \{1, 2\}$ of a potential cluster. Each of these individuals has to choose how to allocate 1 unit of time. Within the potential cluster, core member 1 can collaborate with periphery member 1 while core member 2 can collaborate with periphery member 2. Alternatively each individual can choose to collaborate with an external member to the cluster. The value of such outside collaboration is given by v_c for core member c, drawn from distribution f and v_p for periphery member p, drawn from distribution g. The value of the internal collaboration is

$$v_{cp} = G^{\alpha} V,$$

where V can be normalized to 1 and $G = g_1 + g_2$ is the total amount of public goods provided by core members 1 and 2. These public goods correspond to organizing seminars, training of PhD, etc. The cost for individual i of providing this public good is assumed quadratic $c(g_i) = cg_i^2$.

The timing is the following

1. Nature chooses whether the cluster is selected or not.
2. Core members simultaneously choose how much public goods (g_1, g_2) to provide.
3. The value of outside collaborations is drawn from f and g and core and periphery members choose what collaborations to form.

The utility function of core members is $v - cg_i^2$ and v for periphery members were v is the value of the realized collaboration.

We assume that the formation of a cluster reduces the cost of providing the public good by β. This could reflect a decrease in the cost of organizing activities, or alternatively and

possibly more importantly, could capture an increase in the visibility of the production of public goods and the reputational benefits attached to it.

Results

Without loss of generality, we normalize distribution g to be degenerate with mass on 0. This implies that periphery members will always accept the collaboration with the core member.

The core member decides to collaborate with the periphery partner if and only if

$$G^{\alpha} \geq v_c.$$

The expected number of collaborations internal to the research cluster is given by $2F\left((g_i + g_j)^{\alpha}\right)$.

In stage 2 we solve for the NE of the public good provision game. Fixing the contribution of the other member to g_j, the expected utility of player i if she contributes g_i is given by

$$(g_i + g_j)^{\alpha} F\left((g_i + g_j)^{\alpha}\right) + \int_{(g_i+g_j)^{\alpha}}^{+\infty} v f(v) dv - c g_i^2.$$

Best responses are thus given by

$$2c g_i^* = \alpha G^{\alpha-1} F(G^{\alpha}), \forall i = 1, 2,$$

so that $g_1^* = g_2^* = g^*$ and solves

$$c\ (2g^*)^{2-\alpha} = \alpha F((2g^*)^{\alpha}). \tag{A.1}$$

Thus, if a cluster is formed and c decreases, Equation (A.1) implies that the provision of public good $(2g^*)$ increases and as a consequence, the collaborations within the cluster increase.

The results are summarized in the following Proposition:

Proposition 1 *In equilibrium, each core member of the research cluster provides a level of public good* g^* *such that* $c\ (2g^*)^{2-\alpha} = \alpha F((2g^*)^{\alpha})$ *and the expected number of collaborations within the cluster is given by* $2F\left((2g^*)^{\alpha}\right)$ *and is decreasing in* c.

Interviews of managers of LabEx

We joined a team writing a report for the French government on this policy. Part of the work involved conducting a number of interviews. One of us personally participated in these interviews. The list of people interviewed includes managers in charge of this policy at the top of the public administration (as mentioned above), members of the international committee, as well as 12 researchers (most are the coordinators of the clusters, often interviewed together with their administrative directors).

We explicitly added questions on collaborations in all interviews. Interviews were operated in semi-directed framework, around the 4 following items:

1. Origins of the project: the interviewer was meant to find out the level of prior collaborations between partners.

2. Organization: how was the funding divided, what was the governance?

3. Output: what was the output in terms of publications, tech transfer and collaboration

4. Lessons learned: the respondents were asked to evaluate the strength and weaknesses of the funding instrument.

At the end of the interview, the respondents were asked to give 3 words characterizing the funding scheme. We put this list in the Appendix to the letter, as we find it quite informative.

Overall 10 distinct clusters were approached, 1 in maths, 1 in social sciences, 3 engineering, 2 in biology, 1 in genetics, 2 in physics. Overall 12 people were interviewed.

To sum up the data collected on those clusters, two types of cases emerge. On the one hand LabEx (6 of them) that had no prior experience of collaborations and who highlight the impact of the funding schemes on building collaborations. Here are a few revealing quotes:

- "The project led to real collaborations between the two groups, who barely knew each other before (...) this led to co-publications"

- "Without the LabEx, we would never have been able to build this common structure"
- "We joined people who did not use to collaborate around a common project (...) build something that did not exist before.
- "The LabEx is a tool that allowed interactions between people who were not collaborating before; it favored new collaborations between LabEx members."

On the other hand, there were projects (4 of them) where the members were already working together. In this case, all highlight that it allowed them to strengthen the collaborations.

- "With the LabEx, we reinforced the collaborations between researchers of these two labs, as well as collaborations with other national and international actors"
- "We were already organized as a community. The LabEx allowed us to amplify the interdisciplinarity and to structure the community"
- "The LabEx was a success, because the institutions already used to collaborate together. We tried to reinforce them and make them more inclusive".

Finally the interviews revealed the instruments that were used that favored collaborations. Most of them mentioned setting up internal calls for projects that needed to involve at least two partners, and some of them did the same for PhD supervision. Other comments included.

- "to favor collaboration an essential tool was to have postdoc share an office to create synergies"
- "there were implicit rules that structure research: in each group, for decisions / evaluations, we made sure that people from each lab and each research theme were represented"
- "PhD students and post docs were shared between labs. All this happened naturally and did not need to be formalized".

Overall these interviews strongly confirm the idea that the LabEx induced new collaborations between researchers who did not interact before (or consolidated existing one). The reasons appear to involve a mix between explicit tools set up by the LabEx management, but also implicit, natural rules that emerged.