The evaluation of competitive research funding : .an application to French programs

Marianne Lanoë

THÈSE PRÉSENTÉE
POUR OBTENIR LE GRADE DE

DOCTEUR DE L'UNIVERSITÉ DE BORDEAUX

ÉCOLE DOCTORALE ENTREPRISE, ÉCONOMIE, SOCIÉTÉ (ED N°42)
SPÉCIALITÉ : SCIENCES ÉCONOMIQUES

Par **Marianne LANOË**

THE EVALUATION OF COMPETITIVE RESEARCH FUNDING: AN APPLICATION TO FRENCH PROGRAMS

Sous la direction de **Nicolas CARAYOL**, Professeur des Universités

Soutenue le 07 décembre 2018

Membres du jury :

M. **Tanguy BERNARD**,
Professeur, Université de Bordeaux, *président*

Mme. **Inès MACHO-STADLER**,
Professeur, Universitat Autònoma de Barcelona, *rapportrice*

Mme. **Paula E. STEPHAN**,
Professeur, Georgia State University, *rapportrice*

Mme. **Bérangère VIRLON**,
Docteur, Agence Nationale de la Recherche, *examinatrice*

M. **Nicolas CARAYOL**,
Professeur, Université de Bordeaux, *directeur de thèse*

EVALUATION DU FINANCEMENT COMPETITIF DE LA RECHERCHE : UNE APPLICATION AUX PROGRAMMES FRANCAIS

Résumé:

Cette thèse a pour objet de proposer une évaluation de deux réformes nationales de l'enseignement supérieur et de la recherche mises en place en France au cours des années 2000. Ces politiques répondent à des objectifs différents mais ont toutes deux impliqué une modification du mécanisme d'allocation des financements de la recherche par l'introduction de nouveaux financements compétitifs. Dans un premier temps nous nous sommes intéressés à la création de l'Agence de financement de la Recherche (ANR) qui généralise le recours aux financements sur projets dans une optique d'efficacité. Le premier chapitre analyse si les chercheurs qui choisissent des question de recherche originales ou travaillent davantage sur des projets interdisciplinaires ont tendance à être systématiquement pénalisés lors de la phase de sélection des projets par les comités d'évaluation. Le second chapitre porte sur l'évaluation de l'impact de l'obtention d'un financement sur projet ANR sur les performances scientifiques ex-post des chercheurs financés. Enfin, le troisième chapitre porte sur le Programme d'Investissement d'Avenir (PIA) lancé en 2010, et a pour objectif d'évaluer la mise en place de la politique Initiative d'Excellence (IdEx), qui vise à modifier le paysage universitaire français en faisant émerger de grandes universités d'excellence à l'échelle mondiale. L'étude porte sur l'évaluation de l'impact ex-post de l'obtention du label IdEx par l'université de Bordeaux et de la création de laboratoires d'excellence (LabEx), sur des critères d'excellence de la recherche.

Mots-clés: Financement sur projet, Evaluation de la recherche, Excellence de la recherche, Politique Publique

---

THE EVALUATION OF COMPETITIVE RESEARCH FUNDING: AN APPLICATION TO FRENCH PROGRAMS

Abstract:

This thesis is intended to evaluate the implementation of two national policies in the 2000s, which aim at reforming the Higher Education and Research system in France. Although they have been established with different purposes, they both modified the funding allocation mechanism by the introduction of more competitive funding. We first focus on the creation of the French funding agency, *Agence Nationale de la Recherche* (ANR), which generalized project-based funding with an efficiency rationale. The first chapter analyzes whether researchers who address more original research problems or higher degree of interdisciplinary research tend to be systematically penalized during projects evaluation by the review committees. The second chapter investigates the impact of obtaining an ANR project-based funding on grantees' ex-post scientific performances. Finally, chapter three is aimed at evaluating the implementation of the IdEx policy, called *Initiative d'Excellence*. This program was launched in 2010 in order to modify the French academic landscape and prompt universities located in the same geographic area to form sizable institutions of excellence and able to compete worldwide. We investigate the ex-post impact of the University of Bordeaux Idex and the creation of research clusters of excellence, on research excellence criteria.

Keywords: Project-based funding, Research evaluation, Research Excellence, Public policy

JEL classification: D04, O38

Groupe de Recherche en Économie Théorique et Appliquée (GREThA)

UMR CNRS 5113 - Avenue Léon Duguit - 33608 Pessac - France

Résumé de la thèse en Français

L'objet d'étude de la thèse porte sur le système de la recherche publique en France. Ce pays représente un cas d'étude intéressant puisque la mise en place de plusieurs réformes dans les années 2000 a durablement modifié le paysage de l'enseignement supérieur et de la recherche. Ces politiques publiques ont modifié les mécanismes d'allocation des financements de la recherche et des pratiques d'évaluation, ce qui a introduit une plus grande compétition et distinction entre individus, laboratoires ou encore entre institutions.

Quatre réformes ont joué un rôle majeur dans l'évolution de l'organisation et de l'orientation de la recherche : Premièrement, l'Agence Nationale de la Recherche (ANR) a été créée par le gouvernement en 2005, dans le but de centraliser et généraliser les financements sur projets au niveau national. Ainsi, la création de l'ANR a participé à la modification du mécanisme d'allocation des financements en donnant une place plus importante aux financements contractuels compétitifs. Ensuite, la création de l'Agence d'évaluation de la recherche et de l'enseignement supérieur (AERES) a été effective en 2007. L'évaluation des laboratoires de recherche étant déjà pratiquée avant la création de l'agence (avec par exemple l'évaluation des "unités mixtes de recherche" par le CNRS), l'originalité de cette mesure provient de la divulgation publique des résultats de l'évaluation qui étaient précédemment gardés privé. Aussi cette même année, la loi relative aux libertés et responsabilités des universités (LRU) à été adoptée, et confère plus d'autonomie au universités. Enfin, compte tenu des performances relativement décevantes des universités françaises dans les classements internationaux, une séries de mesures a été lancée, parmi lesquelles les Pôles de Recherche et d'Enseignement Supérieur (PRES) ou encore la mise en place de Communauté d'Universités et Etablissements (COMUE), avec pour objectif commun d'inciter les universités et organismes de recherche situés dans une même zone géographique à se rassembler de manière à faire émerger de grandes institutions multidisciplinaires plus visibles. Le programme IdEx, lancé en 2011, est la plus importante de ces mesures et consiste en l'allocation d'importants montants de financements à un petit nombre d'universités préalablement fusionnées, avec pour ambition de les transformer en leader national mais aussi et surtout de manière à être capable de concurrencer les universités à l'international.

L'objectif de cette thèse est de proposer une évaluation de l'efficacité de deux de ces réformes en estimant notamment leur impact, et de souligner les caractéristiques particulières des différents mécanismes d'allocation des financements qui leur sont associées. La thèse est organisée de la manière suivante : Les deux premiers chapitres se focalisent sur la création de l'ANR. Dans le Chapitre 1, le mécanisme d'allocation des financements est examiné et on cherche à savoir qui sont les chercheur qui participent aux programmes et qui sont ceux qui sont financés. Le Chapitre 2 porte sur l'évaluation de l'impact des financements sur projets de l'ANR sur des mesures de performances scientifiques des financés. Enfin, le Chapitre 3 concerne une évaluation préliminaire du programme Idex, dont le but est de créer des université d'Excellence. L'étude a pour but d'évaluer l'impact ex-post de la mise en place de la politique sur des indicateurs bibliométriques utilisés comme proxies de l'excellence de la recherche pour les chercheurs ciblés.

Chapitre 1 :

L'objectif du premier chapitre est d'examiner si les agences de financement se comportent différemment vis à vis des participants aux programmes dont les travaux de recherche abordent des questions plus nouvelles ou sont associés à un plus fort degré d'interdisciplinarité. Nous définissons ces deux types de recherche comme éléments de la recherche non conventionnelle. Cette idée provient d'une inquiétude grandissante dans la communauté scientifique concernant l'aversion au risque des comités de sélection vis à vis du financement de projets. Si cela est vérifié, la recherche originale ou interdisciplinaire, qui ont toutes deux des chances de produire de fortes avancées de la frontière des connaissances, mais sont aussi associées à une plus forte incertitude des résultats de recherche, pourraient être plus systématiquement pénalisées par les comités de sélection. Etant donné qu'il nous est impossible d'observer directement les caractéristiques spécifiques des projets soumis, nous testons une hypothèse alternative portant sur les caractéristiques de la recherche menée au cours des années précédant les programmes de l'ANR. Nous soupçonnons que si les participants aux programmes, dont la recherche non-conventionnelle peut être perçue comme un facteur de risque par les comités dévaluation, sont associés à une probabilité plus faible de voir leur projet financé, toutes choses égales par ailleurs, alors que ce type de recherche peut être désirable pour faire avancer significativement la frontière des connaissances, cela peut indiquer un biais de sélection négatif dans l'évaluation des projets de ces participants. Afin de tester cette hypothèse, on s'intéresse à la fois à la décision des chercheurs académiques travaillant sur le territoire français de soumettre un projet à l'ANR et à la décision de l'agence/des comités de sélection de retenir et financer un projet. Pour cela, la base de donnée est construite à partir de trois sources : la liste des chercheurs et enseignants chercheurs affiliés à un laboratoire de recherche accrédité par le ministère de l'enseignement supérieur et de la recherche, la liste des chercheurs ayant participé à un programme de l'ANR entre 2005 et 2009 et les publications scientifiques extraites à partir de la base Web of Science. Nous estimons un modèle probit avec une correction du biais d'échantillonnage de Heckman de manière à contrôler l'effet du biais de sélection sur les résultats. Dans un premier temps, nos résultats suggèrent que les chercheurs qui abordent des questions de recherche relativement plus originales ou entreprennent de la recherche avec un plus fort degré d'interdisciplinarité ont tendance à davantage participer aux appels à projets. Cependant, ces mêmes chercheurs sont associés à une probabilité conditionnelle plus faible de voir leur projet financé par l'agence. Tout semble fonctionner de telle sorte que les chercheurs qui entreprennent de la recherche dite « non-conventionnelle » surestiment leurs chances de voir leurs projets sélectionnés et financés par l'agence. En effet, ils ont plus de chance de participer à un appel à projet, à qualité de la recherche et autre facteurs égaux par ailleurs, mais les comités d'évaluation ne semblent pas évaluer le projet de la même manière, puisque ces chercheurs ont moins de chance d'obtenir un financement de l'agence. Dans ce contexte d'asymétrie d'information, on peut supposer que les participants ont davantage de connaissance sur leur projet que les comités. En d'autres termes les participants ont une information privée qui n'est pas observable par les comités, ce qui pourrait expliquer cette différence dans la valeur donnée au projet. Ainsi, il est probable que ces chercheurs valorisent davantage leur chances d'obtenir un financement par l'agence que les autre chercheurs. Ensuite, deux types de projets sont définis par l'agence; les programmes thématiques dont le domaine de recherche est choisi par l'agence (en concertation avec des experts de la communauté scientifique, des représentants d'organismes de recherche ainsi que des responsables R&D de grandes entreprises) et non thématiques, qui sont des programmes plus neutres ouverts à tous les champs disciplinaires. Quand on différencie les chercheurs selon le type de programme auxquels ils participent, on trouve que les chercheurs qui entreprennent de la recherche associée à un plus fort degré d'interdisciplinarité ont plus de chance de participer aux deux types de programmes, tandis qu'uniquement les programmes thématiques attirent significativement plus les chercheurs qui traitent de problèmes plus nouveaux. Ce résultats suggère que les thèmes prioritaires ciblés par l'agence pour les programmes thématiques sont bien définis. Néanmoins, la recherche non conventionnelle est toujours pénalisée dans le processus de sélection des programmes thématiques, alors que le biais négatif concerne uniquement la recherche interdisciplinaire pour les programmes non-thématiques.

Chapitre 2 :

Dans le second chapitre est étudié l'impact du financement sur projet de l'ANR sur diverses mesures de production scientifique des financés. Étant donné que l'agence alloue des financements par le biais de programmes thématiques et non thématiques, nous examinons également quel type de programmes est le plus efficace. L'étude est basée sur la même base de données que dans le Chapitre 1. Les financements n'étant pas distribué de manière aléatoire entre participants aux programmes et étant donné que de nombreux facteurs influencent probablement à la fois la probabilité des chercheurs d'être financés et leur performances scientifiques, on estime l'impact par la combinaison de la technique d'appariement/pondération des contrôles et de la méthode de doubles différences de manière à contrôler pour les différences observables entre participants et les différences constantes dans le temps. Dans un premier temps, plusieurs groupes de contrôles ont été constitués de manière à contrôler pour les facteurs de confusion observables. Le but de cette approche est de rendre l'allocation des financements aléatoire entre les chercheurs financés et le groupe de contrôle. La construction des groupes de contrôles est basée sur l'estimation du score de propension, qui représente la probabilité d'un individu d'obtenir un financement (conditionnellement à un ensemble de facteurs). On ne retient ensuite que le groupe de contrôle le plus similaire au groupe financé, c'est à dire celui qui vérifie au mieux les tests d'équilibrage (équilibrage de l'ensemble des variables utilisées pour estimer les scores de propension entre les deux groupes) et un test de sentier parallèle, de manière à s'assurer que l'évolution des performances scientifiques est parallèle entre les groupes de traités et de contrôles sur les années précédant les appels à projets. On trouve que l'obtention d'un financement ANR augmente assez fortement et significativement le nombre de citations reçu, dans une ampleur de 15%. D'autres résultats suggèrent que tandis que l'impact sur les citations est positif et significatif pour les financés des programmes thématiques, les chercheurs qui obtiennent un financement des programmes non-thématiques bénéficient d'un effet plus fort. En effet, les programmes non-thématiques semblent toujours associés à de meilleurs résultats que les programmes thématiques pour nos variables d'intérêt sélectionnées. Cela signifie qu'il est plus efficace en ciblant des projets de haute qualité, bien qu'il n'est pas exclus que les programmes thématiques puissent avoir un impact différé. On observe cependant que les programmes thématiques attire les chercheurs qui abordent des questions de recherche plus originales.

Chapitre 3 :

Le troisième chapitre porte sur l'évaluation de la mise en place de la politique IdEx, dont l'objectif est de favoriser l'émergence d'Universités d'Excellence en France. L'analyse est basée sur l'étude de cas de l'Université de Bordeaux, qui a été sélectionnée et fait donc partie les lauréats du programme en 2011. L'étude ne porte pas sur tous les membres de l'université, mais se concentre sur des groupes de chercheurs qui sont plus spécifiquement ciblés par le programme étant donné qu'ils sont également impliqués dans un autre volet de la politique qui cible des clusters de recherche, les LabEx. Ces derniers sont définis par le regroupement de laboratoires de recherche ou d'équipes de travail d'un laboratoire. Notre objectif est d'estimer si la mise en place de cette politique ambitieuse a eu un impact ex-post sur des mesures d'excellence de la production scientifique des chercheurs appartenant aux LabEx. De manière à constituer la base de données, l'Université de Bordeaux nous a transmis la liste du personnel impliqué dans chacun de ses LabEx, ce qui nous a permis de les identifier dans la liste des chercheurs et enseignants chercheurs en France. Aussi, les autres scientifiques affectés par la politique, c'est à dire ceux rattachés à une autre université sélectionnée dans le cadre du programme IdEx, ou affiliés à un laboratoire de recherche impliqué dans un des autres LabEx ont été identifiés dans la base et supprimés du groupe de contrôle potentiel. La méthodologie a été déterminée de manière à prendre en compte le fait que la sélection a été effectuée au niveau de larges groupes de chercheurs. Nous utilisons une méthode d'appariement optimal qui utilise la structure multi-niveaux des données, dont l'objectif est de reproduire un essai randomisé contrôlé avec appariement par groupe de deux. Cette approche nous permet d'obtenir les meilleurs combinaisons d'appariements deux à deux de laboratoires et d'individus. Dans un premier temps, un groupe de contrôle est construit à partir du large échantillon de chercheurs non affectés par le programme IdEx, et n'appartenant pas à un LabEx. Pour se faire, l'appariement entre le groupe traité et le groupe de contrôle est réalisé sur la base de potentiels facteurs de confusion observés au niveau individuel. Cependant, en plus d'apparier des individus avec des caractéristiques similaires, nous cherchons également à équilibrer les caractéristiques des laboratoires de recherche, qui constituent probablement des facteurs de confusion de l'effet du traitement. Une fois que le groupe de contrôle est constitué, l'effet causal de l'allocation du traitement est estimé selon la méthode de double différence sur des mesures représentant trois dimension de l'Excellence scientifique, mesurées par l'impact de la recherche, sa nouveauté et l'étendue de sa diffusion. Nos résultats préliminaires suggèrent que la politique n'a pas d'effet significatif sur nos proxies de recherche de haute qualité. Tandis que ce résultat préliminaire ne s'applique qu'au cas de l'université de Bordeaux, on observe que seule la moitié du groupe de traité est finalement utilisée dans l'analyse et que ce sous-échantillon n'est plus représentatif de l'échantillon initial de chercheurs rattachés à un LabEx de l'université de Bordeaux. Les chercheurs avec de meilleurs performances ayant été retirés durant l'étape d'appariement par manque de contrôles similaires, il est ainsi possible que l'effet soit en fait sous-estimé.

Remerciements

Je souhaite tout d'abord exprimer mes plus vifs remerciements à mon directeur de thèse, Nicolas Carayol, pour m'avoir encadré et guidé tout au long de d'élaboration de cette thèse. Je tiens à lui adresser toute ma reconnaissance car sa disponibilité, son professionnalisme et sa rigueur scientifique m'ont beaucoup fait évoluer.

Je suis très reconnaissante aux membres du jury, Pr. Inés Macho-Stadler, Pr. Paula Stephan, Dr. Berangère Virlon et Pr. Tanguy Bernard, de m'avoir fait l'honneur de lire ce travail et d'avoir accepté de participer au jury de soutenance.

Je remercie également les personnes qui ont pu me conseiller à différents moments de la thèse, Pascale Roux pour sa bienveillance, Matthieu Clément et Tanguy Bernard pour leur disponibilité, Francesco Lissoni et Valério Sterzi pour leur générosité et gentillesse. Mes remerciement vont également à Oscar Llopis, Diego Useche et Valentina Di Lasio, dont j'ai apprécié les conseils et l'aide apportée.
J'adresse de chaleureux remerciements à Karine Onfroy et Edouard Kleinpeter pour leur implication dans nos projets communs et leur gentillesse.

Je souhaite également remercier les personnes rencontrées lors des conférences, pour leur gentillesse et leurs conseils, avec une pensée particulière pour Cornelia Lawson.

Je remercie également mes collègues du laboratoires, et plus particulièrement les personnes avec qui j'ai partagé le bureau, Guillaume Pastureau, Emmanuel Lorenzon, Laurent Berger, Elodie Carpentier, François Maublanc et Adil El Aichouchi.

Merci à mes amis pour leurs encouragements, particulièrement à Laurent, Christophe, Huihui, Martine, Alain, Thelma, Julien, Manon et Anna.

Un grand merci à ma famille et mes parents pour leur confiance et leur soutien sans faille, même de loin. Mes derniers remerciements s'adressent à Fabien pour ses encouragements et sa compréhension.

Contents

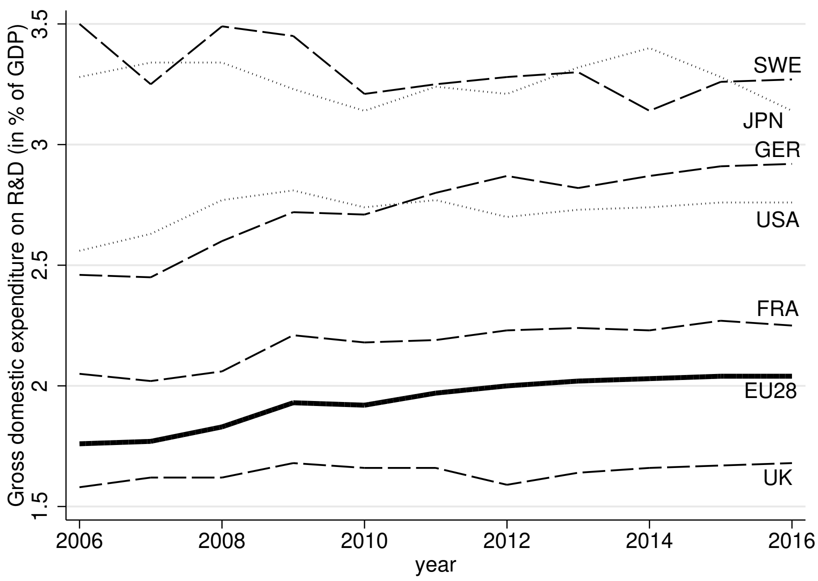
Introduction

One of the Lisbon Strategy objectives, established during the European Council in 2000, was to make European Union "the most competitive and dynamic knowledge economy in the word" in the next ten years. This ambitious goal achievement should have been enabled by an increased investment in research, in order to stimulate knowledge creation and innovation, which in turn should have improved technological progress and economic growth. Thus, a specific objective was set up during the European Council in 2002, which stated that members of the European Union should spend approximately 3% [1] of their Gross Domestic Product (GDP) on research and development (R&D) by 2010. As displayed in Figure 1, this goal was not reached by the group of 28 European Union (EU) countries at the end of the last decade, nor by most of the EU countries alone, that is why its pursue has been renewed in the framework of the Europe 2020 Strategy. However, with an R&D intensity for EU countries slightly above 2% in 2016, this goal is still far from being achieved.

Arrow (1972) and Nelson (1959) were interested in the nature of knowledge, which has the characteristics of a public good. The non-rivalry and non-excludability properties imply it is difficult to prevent people to use a piece of knowledge once disclosed. Therefore, the appropriation of the results of basic research along with its social returns is complicated. Moreover, the market does not give enough incentives for basic research to be funded and performed in enterprises, thus this would lead to a 'market failure', with an under-investment in basic research compared to its socially optimal level. This thus justifies the traditional involvement of the State in the support of basic research performed in public research laboratories and universities.

[1]The administrations would bear one-third of the spending effort, when the private sector would be in charge of the rest.

Figure 1 – Gross domestic expenditure (GERD) on R&D for several EU countries, UK, US and Japan (in % of GDP).



Source: Eurostat.

As presented by Dasgupta & David (1994), the institutional organization of Science is based on a self governance structure which is rather socially efficient. Indeed, the reward system being conditioned on the priority rule, research efforts are encouraged in order to be the first to discover. Then the publicly disclosure of the results is used for claiming ownership of the discovery which allows to establish the researchers' reputation and to receive rewards in terms of peers-recognition, subsequent funding, promotions, etc (Merton 1957, Stephan 1996). The reward system is then based on research achievement rather than the level of efforts performed that are complicated to monitor, and allows then to make up for the appropriation of basic research results issue. Thus, knowledge creation is encouraged by non-market-based incentives associated to the reputation-based reward system in Science.

Though the institutional organization of Science produces incentives to perform research, high level of financial investments are also required, considering research is costly. Not only in terms of researchers salaries but the cost of equipment and research materials can require high level of funding. This concerns many research field and ultimate examples may be found in astronomy with the building of large telescopes or in biomedical sciences such as with machines for the genome sequencing or simply with the purchase of genetically modified mice and their care (Stephan 2012).

Governments of European and other developed countries have become more interventionist in developing science policy initiatives from the mid-70s, in such a way as to respond to a concern of efficiency and accountability. Thus, an evaluation culture has gradually been established, with assessments taking place at the institutional and national level in additional to the individual or team level (for evaluation by funding agencies for instance). Each countries has developed their own funding allocation system and institutional assessment approach based on different practices. For instance in The Netherlands, policy-makers has tended to change the direction of research towards issues of societal relevance, and implemented an institutional evaluation system in order to assess the quality of their organizations. However, the evaluation was performed by the university itself and was rather used to improve the research strategy of the organization than to determine the distribution of funding (van Steen & Eijffinger 1998). Unlike other countries, the UK early implemented a performance-based system in the mid-80s, based on an advanced institutional evaluation system, the “Research Assessment Exercise” (RAE), in order to assess research units of universities on the basis of their scientific publication activity. The ranking was then used to define the allocation of funds from the Higher Education Funding Council (Geuna & Martin 2003).

This thesis focuses on the public research system in France. This country is an interesting subject of study since the recent implementations of several reforms during the 2000s has durably modified the higher education and research landscape. These policies introduced some changes in the funding allocation mechanism and evaluation practices, which implied more competition and distinction at the individual, laboratory and institution levels.

Three reforms have played a major role in the changes of the organisation and orientation of research: First, in 2005 was created by the government the French funding agency, *Agence nationale de la recherche*, which aim was to centralize and generalize project-based fundings at the national level.[2] Thus, the setting-up of the agency participated in the modification of the funding allocation mechanism and gave a more prominent role to competitive grants. As shown in Table 1, the nature of public research funding has evolved in France considering the increase of the share of contractual fundings in the total funding of Universities and public research organizations (PRO).

Secondly, the creation of a French evaluation agency, *Agence d'évaluation de la recherche et de l'enseignement supérieur* (AERES),[3] was initiated in 2007. The evaluation of research laboratories was not new, because it was already performed before the creation of the agency (for instance the CNRS assessed its "mixed units"[4]). The originality rests upon the public disclosure of evaluation results along with a grading of research laboratories, which were hitherto kept private (Musselin 2017). This also led to increase competition, and could have been used as a basis for allocation decision making.[5]

Finally, as regards the disappointing performances of French universities in international rankings, a series of measures associated to a selective participation of institutions have been launched by the government.[6] The measures common objective was to prompt universities located in the same geographic area to gather together in order to make sizable interdisciplinary institutions emerge,[7] which would be able to compete at the international level. Launched in 2011, the IdEx program[8] is the most important of these measures, and consists in the allocation of substantial funding amounts to a small number of selected merged universities, in order to convert them into national leaders and international competitors.

[2] Project-based fundings already existed before, but their allocation was managed either by public research organizations (such as CNRS) or by the ministries.

[3] The Evaluation Agency becomes *Haut Conseil de l'évaluation de la recherche et de l'enseignement supérieur* (HCERES) in 2014.

[4] "Mixed units" are joint units, composed of CNRS researchers and professors working in an university.

[5] For instance, since 2008, the AERES evaluation was used by the ministry, even though faintly, in order to allocate endowments.(Musselin 2017)

[6] Among these measures which aim at gathering institutions, we can mentioned the *pôles de recherche et d'enseignement supérieur* (PRES) launched in 2006 or the *Communauté d'universités et établissements* (COMUE) set up in 2013.

[7] The structure of French universities was based on a mono-disciplinary rationale so far.

[8] The IdEx program is involved in a larger program called "Investissement d'Avenir", which is a massive investment program, a part of whom is especially dedicated to fund public institutions, groups of research laboratories and the building of research equipment.

Table 1 – Evolution of the funding amount according to the source of funds for Universities and Public Research Organizations (PRO)

	2005	2007	2009	2011	2013	2015
Universities						
Block funding	3,427	3,514	4,168	4,746	4,857	4,758
Contractual funding	692	945	1,206	985	1,267	1,458
PRO						
Block funding	6,147	6,726	7,183	6,395	6,608	6,371
Contractual funding	1,621	1,671	2,215	2,709	2,477	2,406
Total contract. fund./Total block fund.	0.23	0.24	0.28	0.33	0.33	0.34
Total contract. fund./Total funding	0.17	0.17	0.20	0.22	0.23	0.22

Source: MESR-DGESIP/DGRI-SIES, L'état de l'enseignement supérieur et de la recherche reports (2007-2017).

Notes: Amounts are in million euros. In 2010, there is a shift in the administration survey methodology to improve information quality, which causes a substantial downwards correction of the amount of spending (from 2010).

Block funding represents the budget allocated by the ministries to universities and PROs (including salaries, operating cost, etc). Contractual funding includes project-based funding (national, such as ANR grants, European and international grants), funding from engagement with enterprises, etc.

PRO stands for public research organizations, and includes all the mission-oriented public entities, such as CNRS, INSERM, INRA, IRD, IRSTEA, IFREMER, etc.

The total amount of funding is not reported, but it also includes the own resources of Universities and PROs.

The aim of this thesis is to offer an evaluation of the main policies recently implemented in France, in terms of efficiency and ex-post impact, and to emphasize the specific features of the different funding allocation mechanisms associated to these policies.

The thesis is organized as follows: The first two chapters focus on the creation of the French funding agency. In Chapter 1, we investigate the grant allocation mechanism and mainly whether the agency behaves differently towards applicants who performed unconventional research, as defined by more novel research or by more interdisciplinary research, in the recent past. Chapter 2 focuses on the impact evaluation of the allocation of fundings to projects by the agency on ex-post measures of grantees' scientific performances. Finally, Chapter 3 concerns a preliminary evaluation of the IdEx program, which aims to make Universities of excellence emerge, in which we estimate the ex-post impact of the policy implementation on research excellence bibliometric indicators of the targeted researchers.

Chapter 1 :

The goal of the first chapter is to examine whether funding agencies behaves differently towards applicants who carry out novel research or interdisciplinary research, that we both gather into the term unconventional research. This idea relies on a growing concern in the scientific community, which claims that peer-review committees would be risk-adverse when they decide which projects to fund. If so, this behavior could leads to a systematic selection bias toward highly innovative research which is more uncertain while it is likely to produce considerable knowledge advances. Considering we can not directly observe the characteristics of the submitted proposals, we test an alternative hypothesis using the characteristics of the applicants. We suspect that if applicants, whose unconventional research performed could be seen as a risk factor by the evaluation committees, are less likely to have their project funded, other things being equals, while that type of research may be desirable to make the knowledge frontier advance, this evidences a downward bias in the project evaluation of these applicants. To test this hypothesis, we are both interested in the decision of French researcher to apply to an ANR program and the decision of the agency to grant a project or not. We find that researchers who address more novel problems or carry out higher degree of interdisciplinary research, defined as unconventional research, are more likely to apply. However, we find that these researchers are penalized by the evaluation committees and are less likely to be awarded a grant. Indeed, both novelty and interdisciplinarity of research are negatively associated with the likelihood of an application to be successful. Considering different type of programs may follow different objectives when selecting projects to grant, the analysis is then done in differentiating directed from non-directed programs. We find that researchers who perform more novel research or more interdisciplinarity are less likely to be funded in the directed programs, whereas the non-directed programs only penalize researchers who carry out interdisciplinary research. However other results suggest that researchers who apply to the directed programs perform more novel research than those who participate in non-directed programs. Thus, directed programs seem well designed to attract researchers who tackle original research approaches.

Chapter 2 :

In the second chapter, we investigate the impact of projects funding by the French funding agency (ANR) on various scientific outcome measures of the grantees. Since the agency allocates grants through both directed and not directed programs, we also analyze which type of program is the most efficient. Considering fundings are not randomly assigned among applicants and that many factors are likely to influence both the likelihood of researchers to be funded and their scientific performances, we estimate a conditional difference-in-differences model in order to adjust on both observable differences in the attributes of the applicants and time-invariant differences. We first design several potential groups of control in order to adjust for the observed confounding variables. The goal of this approach is to make funding allocation random between the treated and control groups. Each control group construction is based on the estimated propensity score, which represents the probability of the individual to obtain a grant (conditional on a defined set of attributes). We finally only retain the group which is the closest to the group of grantees, namely the group that best verifies both the balancing tests (balance of the set of covariates used to estimate the propensity score between the two groups) and a placebo parallel path test (to ensure that performance path are parallel between the granted and control groups before the funding assignment). We find that obtaining an ANR grant increases substantially and significantly the number of ex-post citations received. Other results suggest that whereas the impact on citations is positive and significant for grantees of the directed programs, researchers who obtain a grant from the non directed program benefit from an even larger effect. In fact, the non directed program always seems to perform better than the directed programs for our outcome covariates, that suggests it is more efficient in targeting high quality projects, even though we do not exclude that directed programs may have delayed impact. We observe however that directed programs attract researchers that carry out more novel research.

Chapter 3 :

The third chapter is related to the evaluation of the IdEx policy implementation which aims at promoting the emergence of Universities of excellence in France. For this purpose, our study is based on the University of Bordeaux, which has succeeded the selection process in 2011. We precisely focus on groups of researchers who are more specifically targeted by the program since they are also involved in another component of the policy, which this time targets research cluster, defined as a gathering of research laboratories or work teams of laboratories, and not the whole University. Our goal is to estimate whether the policy has an ex-post impact on the scientific production of this group of researchers, as measured by three dimensions of scientific excellence defined as research impact, novelty and diffusion. In a first step, we built a control group composed of researchers drawn from the large sample of researchers neither affected by the main IdEx policy program, or by the research cluster component. We select individual-level factors for the matching but also seek to balance laboratory-level covariates, which are both likely to confound the treatment effect. We use an optimal matching method which makes use of the multilevel structure of the data and finally produces the best combinations of matched pairs of laboratories and treated/control individuals. We then estimate the causal effect of assignment to treatment, relying on the the difference-in-differences method in order to adjust for remaining time-invariant differences between the treated and the matched groups. Our results suggest that the policy does not have any significant effect on our ex-post scientific excellence bibliometric indicators. Whereas this preliminary result only applies to the University of Bordeaux case, we observe that the halved treated sample used in the analysis is no longer representative of the whole sample of researchers belonging to a research cluster in Bordeaux besides. Treated researchers with higher performances being removed, therefore it is possible that the effect is underestimated.

Chapter 1

Do public funding agencies support novel research? A study of the competitive grant allocation in France.

Introduction

Public funding for research is increasingly based on a competitive allocation mechanism. This mode of allocation has been implemented a long time ago in the USA by funding agencies, and in many European countries more recently (Geuna 2001).

The basic mission shared by public funding agencies is to promote research excellence towards advancing the knowledge frontier. One way to achieve this goal is to provide resources to talented researchers when they engage in promising research areas or conduct research in unexplored directions that likely lead to innovative results. This goal is for instance put forward by the European Research Council, which 2017 budget amounts to 1.8 billion euros, as it aims to “support excellent investigators and their research teams to pursue ground-breaking, high-gain/high-risk research” with the intention of “pushing forward the frontiers of knowledge” (ERC 2017, p.14). Highly innovative research can lead to scientific breakthroughs. It is characterized by a higher risk of failure but is also likely to lead to high impact results (Uzzi et al. 2013, Carayol et al. 2018), which are themselves further expected to lead to new high impact results (Wang et al. 2017).

However, the scientific community is increasingly concerned with a systematic negative bias toward research novelty in peer-review evaluation (Petsko 2012, Nicholson & Ioannidis 2012, Stephan 2012). Reviewers' risk-aversion would tend to give an advantage to seemingly safer conventional research ideas rather than promoting original approaches whose results are more uncertain. In this way, the choice between, on one hand, potential “high gain/high risk” projects and, on the other hand, more conformist proposals based on the exploitation of existing knowledge usually ends in favor of the latter. It is likely that funding agencies become aware of this problem, as it is signaled by the launch of programs entirely dedicated to fund explorative research. For instance, one of the program of the National Institutes of Health (NIH), which is the largest funder of biomedical research in the world, is dedicated to encourage extremely creative researchers to conceive highly innovative research frameworks with potential broad impact. The Transformative Research Award launched in 2009 is one of the funding opportunities which fall within this program.

In this chapter, we investigate whether funding agencies behaves differently towards researchers who performed unconventional research in the recent past. We define non conventional research in two ways, by novel research, and by interdisciplinary research. Both of these types of research are associated with potentially high innovative results, but are also considered as more risky or uncertain. As regards the growing concern about risk-aversion of panel committees during the projects evaluation process, and a possible negative bias towards original projects, we hypothesize that if researchers who carry out novel research or interdisciplinary research are less likely to be funded, it may signal a downward bias associated with the evaluation of their projects. We first observe who have more incentives to apply, according to their attributes and past research performances, in order to investigate the behavior of researchers who perform non conventional research. And then, our goal is to evidence whether those researchers are penalized by the evaluation committees during the grant process. We also investigate whether different programs designs, which differ in terms of their objectives or how they choose the review committees, lead to the same findings. Finally, we examine whether the diversity of projects teams composition, often associated to potentially innovative projects, as measured for instance by the degree of multidisciplinarity in the team, is also penalized during the projects evaluation process.

Our study is based on the French funding agency *Agence Nationale de la Recherche* (ANR). This national public organism has been created in 2005 on the French government's initiative. Its objective was to centralize the majority of contractual competitive grants in France. This type of funding already existed before the creation of the agency, although allocated amounts were lower. These contractual grants were managed by the public research organizations (PROs), such as the National Center for Scientific Research (CNRS), which is the largest European public research organization, or the National Institute of Health and Medical Research (INSERM), which decided independently how to allocate these funds in their own departments. The total budget of the agency was initially intended to quickly exceed one billion euros. However, the budget increased from 700 million euros to 850 million euros between 2005 and 2008, and then gradually decreased until 2015, whereas the number of applications has continuously raised. Since 2010, the scope of responsibility of the agency has increased with the launch of new programs directed to allocate competitive funds to consortium of universities and to groups of laboratories to promote research excellence. The agency is responsible for the formulation of the calls for proposals in interaction with the government. The ANR is also in charge of the selection of awarded projects which is performed by peer-review panels involving international experts.

The missions announced by the ANR are close to those of the other funding agencies. The foremost concern is to stimulate research excellence through funding. Our study focuses mainly on two of its objectives, described as the intention to foster creativity and emerging areas, that we represent by novel research, and the wish to encourage interaction between research fields,[1] that we sum up as interdisciplinary research.[2] Our goal is to assess whether the funding agency conforms to its commitments to promote novel approaches and interdisciplinary research during the grant process. Although we cannot observe directly the quality and originality of the submitted projects, we study whether research based on non conventional approaches, namely original research or interdisciplinary research, that we measure from the past scientific production of the investigators of the projects, are more likely to be funded by the agency. While we only refer to project-based funding programs, we also investigate whether the different program designs target distinct categories of researchers. We are indeed able to distinguish between directed programs, whose research areas are predetermined by the agency, and more standard non-directed programs. The themes of the directed programs are defined within the scope of identified key societal challenges and thus should be especially intended for creative researchers who address novel research approaches.

Our study is based on a database which consists of approximately 30,000 French researchers and professors, among which about one third applies at least once to the ANR between 2005 and 2009. This structure of our data allows us to model the grant process using a Heckman probit model in order to correct for selection bias. The only use of the subsample of applicants to investigate which determinants influence the funding decision would indeed bias the estimated results given that the decision to apply does not follow a random process. Researchers relatively more advanced in their career, those

[1] http://www.agence-nationale-recherche.fr/missions-et-organisation/missions/ consulted the 18/10/2018.

[2] We employ the term "interdisciplinary research" in the broad sense of multiple discipline research.

affiliated to top-ranked laboratories and those characterized by better quality of past performances are systematically more likely to apply. This evidences self-selection.

Our first results show that, controlling for the quality of past performances, researchers who explore novel approaches and perform higher degree of interdisciplinary research are more likely to submit a project to the agency. However, these researchers are also those who are penalized by the peer-review committees. We find that both novelty and interdisciplinarity of research are negatively and significantly associated with the conditional probability of the application to be successful. We also investigate theses effects according to the type of ANR programs. We find that while the grant process of directed programs, based on the agency requested research announcements, are affected by both of these negative bias, the non-directed programs, defined as unsolicited research projects, seem to put only interdisciplinary research at a disadvantage. However other results suggest that participants in the directed programs perform more novel research than applicants in non-directed programs. The former programs seem then well designed to attract creative researchers with original research approaches. However, programs that do not specify any direction for the definition of research subjects would mainly attract researchers that carry out relatively more conventional research.

The remainder of the chapter is structured as follow. We discuss in the first section I how non-conventional research is evaluated by peer-review in the literature. In the second section II we present the French funding agency and its objectives. Then we describe the sources of our database along with descriptive statistics in Section III. In the fourth section IV we develop the methodological approach and we finally present the results in Section V.

I Funding non-conventional research

Novel research can be defined as the creation of a new piece of knowlegde (a new instrument for instance) but refers more often to the new combination of multiple pieces of existing knowlegde. Jacob (1977) connects this creation process to the new understanding of distinct components, which have been seen unrelated but makes a new idea emerge when "looking at objects from a different angle". Novel research and interdisciplinary research are connected concepts. The definition of novelty as the recombination of distinct ideas easily reveal an overlap area with interdisciplinary research. Novel research does not necessarily combine ideas from distinct scientific field, nor interdisciplinarity necessarily brings distinct ideas together in an original way.

When they address complex questions, researchers may face a problem of missing knowledge or instrument in their own scientific field whereas the needed piece of information has already been developed in another field. The establishment of connections between different bodies of knowledge, in crossing the own's discipline frontier, appear to be useful to solve some issues. Interdisciplinary research is not performed per se, but represents a tool to achieve a complex goal. Those interdisciplinary projects are also associated with potential high impact results and some have lead to major breakthroughs. In the History of Science, this connection between different disciplines has led to the creation of new research lines but also to the start of new research fields, such as the emergence of Quantum Mechanics initiated from the theories of distinguished researchers such as M. Planck, A. Einstein and E. Schrödinger, and defined at the frontier of Chemistry and Atomic Physics. However, this type of research is characterized by a higher risk of failure, just as novel research is.

Researchers can find that unconventional research is not enough rewarded, given that peer-review can be more reluctant to validate very original results, such as in revues (Ioannidis et al. 2014) or that orthodox researchers can be defiant to further exploit these new knowledge. For these reasons, it may exist a natural under-investment in novel or interdisciplinary research, which should be encouraged by funding agencies.

Several explanations are suggested in the literature to account for the potential negative tendency towards the funding of non-conventional projects. Original research is more complex to evaluate since it implies knowledge creation or recombination of possible distant knowledge. The peer-review committees in charge of the projects assessment may thus be more uncertain about the expected quality of the results and more incline to reject the projects. It is also argued that interdisciplinary research can be penalized compared to monodisciplinary research. Science is organized in disciplines, which are delimited by their own frontiers. So peer-review might underestimate interdisciplinary research given that evaluators are influenced by their own knowledge and beliefs and that they may tend to evaluate interdisciplinary research with their own disciplinary standards. It has been found that the assessment of interdisciplinary research is biased, even when evaluated with an interdisciplinary panel (Mallard et al. 2009). According to Laudel (2006), "Interdisciplinary peer review of grant proposal is particularly problematic because it is necessary to synthesize several disciplinary opinions, and a multidisciplinary panel has difficulties reconciling different perspectives" (Laudel 2006, p.59).

Another argument is that it becomes more difficult to rank the proposals when the number of application get larger. It has been argued that some research teams enroll in an application race to get more fundings (Alberts 2010). The existence of a bias against novelty during the grant process is a more sensitive concern in a context of scare resources and low success rates of applications (Fang et al. 2016). As long as the share of granted projects is acceptable, highly innovative quality projects should probably be awarded a grant even if a worse grade was assigned to these risky projects by reviewers. In this situation, the evaluation process is akin to separating good proposals from the one that have some flaws, some feasibility issues or that are not well designed. While in a situation with very low success rates reviewers need to have enough expertise in the area to differentiate excellent projects from the very good ones even though it is difficult to forecast the potential outcomes of innovative projects. For many funding agencies, the success rate of applications have highly decreased. This is linked to the fact that the budget has remained rather low whereas the application rate has increased. In 2016, the mean success rate of the European Research Council for advanced grant was only 9.6%.[3] Moreover, a share of the budget of some agencies is not only allocated for research but also for different activities such as the building of new facilities (Alberts 2010). This implies a waste of time for many applicants who turn away from their research activity to conceive proposals even though they will not be eventually funded. Given that the furtherance of research in a particular direction may be stopped if the researchers did not get a grant, any failure in the evaluation mechanism may imply some losses for the society if high

[3] http://www.horizon2020.gouv.fr/cid115433/appel-erc-advanced-grant-2016-231-laureats-retenus.html consulted the 20/10/2018.

impact research projects are not funded.

The effects of competitive research funding allocation have been frequently studied in the literature. Some authors deal with the efficiency of competitive funding allocation and estimate the impact of receiving a grant on ex-post measures of scientific performances and find a positive impact on the grantees (Jacob & Lefgren 2011, Gush et al. 2018, Azoulay et al. 2011). The allocation mechanism may also influence how the grantees select their research agenda. Azoulay et al. (2011) compare the impact of two types of program on the ex-post research novelty. They find that the allocation of long term funding to promising researchers by the Howard Hugues Medical Institute outperformes in terms of research novelty the preeminent recipients of a contractual NIH grant. However, Wang et al. (2018) find an opposite conclusion in investigating whether the competitive allocation mechanism reduces innovative research compare to the allocation of block funds. They find that Japanese competitive fundings allow to produce more original research results than conventional block grants, apart from young researchers and women.

These preceding studies suggest that the agencies tend to fund good projects, and show that the allocation mechanism affects the way research is conducted afterwards. However, these results do not give any insights about the reliability of the funding allocation mechanism. This issue also received a great attention in the literature, and sometimes led to divergent conclusions. Some studies show that the priority scores assigned to proposals are positively correlated with some measures of quantity and quality of ex-post publications, showing that the review process performs well in selecting the best proposals. However, most of the other investigations point some difficulties to evaluate accurately the proposals. Park et al. (2015) find that the best NIH rejected projects, finally funded by means of an unexpected extra funding, underperform the projects granted in the first stage. Li & Agha (2015) also find a positive correlation between the score of the proposal and some quality measures of ex-post performances. However, Fang et al. (2016), which use the same database as Li & Agha (2015), find that the ranking of the proposals is not predictive of the ex-post outcomes when they discard the funded proposals having the worst grades. Finally, Bornmann et al. (2010) find that the best rejected proposals are associated with better ex-post performances than granted projects on average, but that this result is not valid for all the scientific fields.

The reliability of the evaluation process performed by expert committees has been criticized for a long time in the literature. The introduction of a bias in the assessment of the proposals quality is likely to modify which projects have finally to be funded. Cole et al. show that the funding of a project depends greatly on who are the appointed reviewers. They find substantial disagreements in the rating of a given proposal by different reviewers and that the variation in the rating of a same proposal by different reviewers exceeds the variation in the final rating of all the proposals (Cole & Simon 1981). From a simulation describing the peer-review process, Day (2015) shows that the introduction of a small bias in the assignment of scores to proposals can prevent a good project to be funded.

Bias is not only introduced according the subjective idea of the reviewer about what type of research proposals deserves to be funded, but may also depend on the expertise of the reviewers. Li (2017) shows that reviewers tend to favor proposals in their own areas, which are also more accurately assessed, or explain that evaluators make just less efforts to support projects outside of their specific research area (Travis & Collins 1991). Despite some attempts to warn researchers about the growing importance that some bibliometric indicators hold in all peer-review processes and their misuse (see Hicks et al. 2015 for the Leiden manifesto for instance), it is common practice for panel committees members to rely on standard past performance metrics of applicants, such as short term impact measures of their publications, in order to evaluate the quality of the projects (Stephan et al. 2017). Some funding agencies also directly request researchers to join information about their publications along with their application. Arora et al. find that short-term past performance measures of the principal investigator influence both the likelihood to receive a funding and the amount of the grant (Arora et al. 2000, Arora & Gambardella 2005). This implies an unfair allocation of funding between researchers, in line with the Mertonian "Matthew effect", whereby the most established researchers have a higher probability to benefit from fundings and then from successive further advantages (Merton 1968, 1988). Moreover, ex-ante performances-based evaluation would not necessarily fund researchers who are likely to perform "high gain/high risk" research because this approach tends to encourage the funding of conventional projects (Geuna & Martin 2003). Moreover researchers whose publications are highly cited do not necessarily mean they performed original research, given that novel research can be associated with delayed recognition, along with a lower likelihood to be both highly cited in the short term and published in high impact factor journals (Stephan et al. 2017). The most cited publications of a researcher may not be the most original ones, whereas some highly cited results are associated with incremental research (Ioannidis et al. 2014). The use of short term metrics also tends to favor the submission of "safe" projects, given that future peer-review evaluations will also refer to the present research results.

Other attributes than the past performance metrics of the investigators may also influence the funding decision, such as their status, or whether they are affiliated to a renowned institution (see Marsh et al. 2008 for a review). Ginther et al. (2011) also find that the ethnicity of the applicants has an influence on the NIH funding decision of projects.

To investigate the relationship between the project novelty and the priority score assigned to the projects, Boudreau et al. (2016) set up an experiment and test whether peer-review committees penalize original research projects. The find that higher scores are assigned to more conventional research proposals compared to novel projects, along with a negative relationship between score and the degree of novelty of the proposal. In a different framework, Banal-Estañol et al. (2018) study whether the grant process of the UK's Engineering and Physical Sciences Research Council is biased against potential high impact projects. Given that team diversity is often connected to innovative research in the literature, they build several measures of structural diversity of the projects teams and find that those composed of members that differ in terms of ability, education or discipline are negatively biased during the grant process and so are less likely to be awarded a grant.

According to the previously mentioned evidences that peer-review committees do not accurately evaluate and select the best proposals, some researchers think that this selection mechanism is not appropriate anymore while low success rates, and suggest more fair alternative selection processes. Ioannidis (2011) suggests several alternatives which include to give a small amount of fund to all the applicants, or to use a random selection of funded projects. Fang & Casadevall (2016) suggest a hybrid mechanism in which review panel would separate in a first step the good proposal from the not well conceived projects. The selection of the grantees would then be done randomly.

II Presentation of the French funding Agency (ANR)

II.1 The agency operating system

The primary objective of the ANR is to fund high quality national research projects, by the support of academic and industrial players. For this purpose, the agency sets up its strategic planning in order to direct the research investigations of stakeholders towards some issues at stake for the upcoming decades. These high priority areas are defined by the ANR, which invites industrial and academic experts, in interaction with the national government in order to anticipate upcoming technological evolution. Their decisions are also influenced by the key challenges identified at the international level, by European policies, by the Grenelle Environment Forum for energy or biological resources issues and by the national research priorities. A variety of directed programs are then designed around these promising areas in order to stimulate French researchers to participate in these technical advances and foster innovation process. Another type of programs launched by the agency consists of more standard non-directed programs opened to all research fields and that allow the researchers to freely defined their project topic before the submission. This type of programs is more competitive than the former, by reason of its wide audience. These two types of programs are designed with different expectations. Some programs are intended to foster exploratory and emerging research (such as the non-directed programs), while others encourage applied research along with industrial development, such as many directed programs in Energy or Environment areas. Programs also differ in their eligibility conditions, some of them only target young researchers while others claim a public/private partnership or a cooperation with an European university to conduct the project. In spite of these differences, we find in most of the texts of the calls for proposals, made available on the ANR website for the different programs, that the agency promote original and innovative projects. Interdisciplinary research are also often mentioned in a favorable manner in the texts. This is particularly noticeable for some highly competitive non-directed programs, for which the importance of originality and multiple disciplines research is emphasizes.

While both type of programs appear to be devised to target researchers who carry out original research approaches and interdisciplinarity, we also examination whether the directed and non-directed programs attract and fund different categories of researchers. Indeed it is likely that the expectations of the panel committees are different as regards the designs of the programs.

II.2 The evaluation process

The evaluation of the projects consists of two levels of peer-review.[4] The first level of evaluation is performed by a panel composed of national and international researchers coming from the public and private sector. Committees are formed in such a way that members share some expertise with the field of the project, so as to be able to assess its quality and suitability. While for the directed programs the committees consist of reviewers from varied disciplines, non-directed programs are evaluated by disciplinary panels.[5] We presume therefore that the evaluation process can potentially be biased given that some disciplinary reviewers may not be able to appreciate the potential impact of interdisciplinary projects. These committees rate the proposals on a scale ranging from "C" to "A". For this purpose, they rely on the comments made by two external experts appointed by the agency in order to assess the quality of a particular project and its feasibility. After this first evaluation, only the subset of projects which obtained the grade of "B" or "A" is sent to the steering committee for the second level of evaluation. This panel committee selects the best proposals to be funded along with an additional list of projects. The decision to select a project is based on the defined primary objectives, the recommendations given by the first review panel, and under the financial constraints of the agency budget. The final decision is taken by the Governing Board, and the selected projects are funded if all the claims have been validated (some negotiations may then be introduced between the ANR and the applicants, as regards the budget requested for

[4]Since then, the agency slightly modified the evaluation process. Henceforth, the applicants are requested to firstly submit a pre-proposal, and if shortlisted, they are then allowed to send the full proposal which is evaluated by committees.

[5]It is stated in the non-directed programs forms that the judgment of an additional peer-review panel with expertise in a different research field can be requested in case of interdisciplinary projects.

instance).

The two appointed experts are in charge of grading the projects according to some evaluation grids which are specific to the specific program. Some of these documents are available on the ANR call for proposals website. They are usually requested to give a grade ranging from 1 to 5 to four to six defined criteria, which may vary from one program to another and over time. In most cases, the criteria involve the assessment of the excellence of the projects, its potential impact and innovative nature. Moreover, the feasibility of the project is also highly valued regarding the expertise of the included partners, their synergy or the consistency of the requested budget according to the project objectives. However, it seems that some evaluation grids are conceived in a way so that the excellence of the principal investigator, based on his past scientific production, along with the reputation of the other co-investigators, are at least valued as much as the relevance of the project.[6] As a common practice shared by funding agencies, ANR also requests a list of some of the most relevant publications of the applicants, besides CV, in order to demonstrate their expertise with respect to the project. We then hypothesize that the quality of the past publications of the investigators may play an influential role in the funding decision.

In the next section, we present our database and some descriptive statistics.

III Research design

In this section, we present how we construct our database and we define the main variables used in the analysis.

III.1 The database

Our database comes from three different sources. Firstly, we use the set of all tenured researchers and faculty members in France, who are affiliated to a French laboratory certified by French Ministry of Higher Education and Research between 2009 and 2012. We have available information about the scientists, such as their surname and first name,

[6] Example of an evaluation grid conceived for a non-directed program of 2006, http://www.agence-nationale-recherche.fr/suivi-bilan/historique-des-appels-a-projets/appel-detail1/programme-blanc-2006/ , consulted the 19/10/2018

their year of birth, their gender, their fine-grained scientific discipline called *section*,[7] and the names of the institution and laboratory where they are working. The initial list consists of 49,225 tenured scientists, from which we remove individuals with incomplete information. This cleaning reduced the base to 48,328 individuals.

Then, we are able to identify from this list those who submitted a proposal in response to a call launched by the French funding agency and those who have been funded. We use an initial base which consists of 67,812 partner×project, and includes the surnames and first names of the scientific investigators, along with some information such as their status, the laboratory and institution where the applicants work, their role in the project (PI or not), the project duration and the value of the grant. After we remove the cases with missing or inaccurate information, and those not affiliated to a public institution, we are left with a subset of 54,852 partner×project. Nearly 32% of them are awarded a grant by the agency for a total amount of 2,4 billion euros, that represents about one quarter of the total costs of these projects (Table 1.1). We then match this list of applicants to our administrative list of researchers and faculty members. We first use an exact matching on the surname and first name of the individuals, and then a fuzzy matching to allow spelling errors in the writing of the surnames. A systematic comparison case by case of the researcher's information completes this second matching to avoid homonymy issues. This approach enables us to identify 10,722 individuals who submitted at least one project between 2005 and 2009, of which 5,786 were at least once selected to be awarded a grant.

In order to build our measures of research novelty and interdisciplinarity, we extract the publications of the scientists from Thomson Reuter ISI Web of Science (WoS) database. The whole set of publications for each individual being gathered on the surname and first name initials, we develop a disambiguation process based on the "seed+expand" approach (Reijnhoudt et al. 2014). The use of this process allows us to detect the papers authored by a homonym researcher. Our approach is based on restrictive conditions in the first "seed" step, followed by a loosened "expand" step. The different steps of the disambiguation

[7] These *sections* also reflect the employer of the researcher, each type of institution (universities or public research organizations such as CNRS, INSERM, INRA, INRIA and IRD) having its own classification.

Table 1.1 – Number of partner×project and number of partner×project funded by year of application for the scientific investigators employed by a public institution

	2005	2006	2007	2008	2009	Total
# of partner×project	5 616	12 881	11 655	9 769	14 931	54 852
# of partner×project funded	3 553	4 188	3 496	3 315	2 890	17 442
total amount of funding	417	511	483	522	470	2 403
Total costs of projects	1 510	2 160	1 910	2 040	1 890	9 510

Amounts are expressed in million euros.

process are presented in Appendix A.[8] This approach allows us to gather documents published from 1999 to 2013 by the researchers and faculty members in our database. At the end of the disambiguation process, we are left with 42,130 unique researchers, among which 9,107 submitted at least one project to the ANR between 2005 and 2009 and 4,895 are awarded at least one grant. Some individuals are then removed from the database for several reasons: first, we delete individuals with a highly common surname who are more likely to retrieve false positive publications. We also remove fine-grained scientific disciplines associated with a very low number of publications retrieved from the WoS. These disciplines refer to almost all researchers from Human and Social Sciences, whose main disclosure activity is not the publication of articles in revues, or these journals are not necessarily included in the WoS data set. Finally our operational database consists of 27,031 researchers and faculty members, among which 9,088 apply at least ones to a call launched by the ANR between 2005 and 2009 and 4,801 have been funded at least once over that period.

The number of partner×project in our final sample is presented in Table 1.2. The number of projects submitted in 2005 is lower compared to the other years, as we already observed for the full sample of applicants employed by a public institution (in Table 1.1). Indeed, 2005 corresponds to the first year of operation of the agency, and while half of the directed programs calls already existed before, the remainder of the specific programs was unfamiliar to the scientific community.[9] The mean success rate is about 30% each year in each program, the rate being higher in 2005 and lower in 2009. This gap is related to the

[8]To have a better understanding of the disambiguation process, see also Chapter 2, Appendix G, p.160

[9]This year is also characterized by a late definition and disclosure of the calls, researchers having a limited time period to prepare their proposals (approximately 48 days only, according to ANR 2005).

difference in the degree of competition introduced by the higher number of applications submitted the last years, which was not absorbed by the very slow increase in the ANR budget over the period.[10] These success rates are slightly higher than those observed in the initial database of applications which were located between 20% and 25 % according to the application year (apart from the year 2005 associated with a higher success rate). Programs calls are also characterized by a high variation in their success rate, according to the specificity of the issues they address. Unlike the non-directed programs, some directed programs with relatively low competition are characterized by a success rate higher than 50%.

In our final sample, the distribution of the applicants into the non-directed programs or one of the seven directed programs is almost balanced over the period (Table 1.3). However the distribution of the applications is heterogeneous among directed programs. The program which supports projects related to the disciplines of Biology and Health attracts the highest number of applicants,[11] that represents 40% of the partner×project in the directed programs. This is also the program which defines the largest number of calls over the period. The directed program which funds Human and Social Sciences (SSH) is characterized by a very low number of participants, however this is mainly due to the removal of many SSH researchers in our sample.

Table 1.2 – Number of submitted projects and number of partner×project according to the application date for the final sample

	2005	2006	2007	2008	2009	total
# submitted projects	1,264	3,214	3,060	2,569	4,025	14,132
# granted projects	859	958	906	779	778	4,280
Success rate (in %)	67.96	29.81	29.61	30.32	19.33	30.29
# partner×project	2,000	4,858	4,591	3,724	5,890	21,063
# granted partner×project	1,331	1,487	1,392	1,219	1,181	6,610

We define three samples of applicants for the analysis. We first consider each of the 9,088 scientific investigators nested in one of the 14,132 project team and treat all these

[10]The ANR budget dedicated to calls for proposals goes from 539 billion euros in 2005 to 650 billion euros in 2009

[11]Life-Science is also the most represented discipline in our sample (see Table 1.6)

Table 1.3 – Distribution of the number of partner×project according to the ANR program and to the application date for our final sample

Programs	2005	2006	2007	2008	2009	Total
Non-directed programs	572	2,196	1,902	1,466	4,240	10,376
Directed programs:	1,428	2,662	2,689	2,258	1,650	10,687
Biology & Health	252	1,252	1,171	998	598	4,271
Renewable Energy & Environment	114	107	116	172	234	743
Ecosystems & Sustainable Development	278	319	404	355	342	1,698
Engineering, Methods & Security	0	0	228	219	87	534
Materials & Information	784	903	0	0	0	1,687
Human and Social Sciences	0	81	90	49	50	270
Information and Communication Sc. & Tech.	0	0	680	465	339	1,484
Total	2,000	4,858	4,591	3,724	5,890	21,063

Notes: The seven last programs are the directed-programs. The low number of applications in the program Human and Social Sciences is due to the removal of many SSH researchers in our sample.

applicants individually. Although the skills of all the members of a team is supposed to matter during the evaluation of the projects, the principal investigator of the project is supposed to carry weight. We then define a second sample composed of the 8,685 leaders of a project team identified as the PI. In these two samples, each researcher is observed once per year (i.e. five times), but multiple times applicants can be observed more often (one observation per partner×project). Nevertheless, researchers that we deemed as potentially inactive, according to their age, can have less than five observations.[12] Finally, it is likely that the evaluation process is not only based on the investigators taken independently, but that the relevance of the make-up of the team is also influential for multi-partner projects. For this purpose, we create a third subsample which does not consider the applicants individually but instead refers to the teams, since each one is represented by one observation. A team is defined as a group composed of at least two scientific investigators associated with a project, that lets us with 5,057 observations. Unlike the two first defined samples, this data subset is not used in the analysis along with the non-participants, given that we do not observe any work teams for the non-applicants.

[12]Researchers whose age is below 24 years old or above 65 years old in a particular year are considered as either too young to work or above the retirement age. In that case, we do not include them in the sample for the corresponding years.

III.2 Construction of the variables

Diverse information associated with the gathered WoS publications allows us to construct some indicators which reflect different dimensions of the research activity in the years preceding the date of the potential application.

We present how we create the variables of interest (novel research and interdisciplinary research), along with the control variables. Two sets of variables are defined, depending on whether we consider the investigators taken individually (i.e. each member of the team) or whether we show an interest in the composition of the teams. All the main measures of scientific performance are based on the articles published in the three-year period preceding the date of the potential application. The variables used in the analysis are described in Table 1.4 and Table 1.5, and some details about their construction are presented in Appendix B.

Individual variables

Different indicators have been developped in the literature in order to capture the novelty of research. Some measures are based on the references cited by the articles in their bibliography. In this case, the novelty indicator may rely on atypical pairwise combinations of journals which publish these references (Uzzi et al. 2013, Lee et al. 2015, Wang et al. 2017), or may be based on the overlap between the sets of references of pairwise combinations of articles from the same knowledge domain (Trapido 2015). Other studies refer to the use of keywords associated with the articles. Azoulay et al. (2011) calculate several measures based on the "Medical Subject Heading" (MeSH) keywords to represent research novelty, such as the age of a keyword or the degree of overlap in MeSH keywords between two sets of articles. The novelty measures used by Boudreau et al. (2016) are also based on MeSH keywords but rely mainly on new pairwise combinations of these keywords. In our study, we use the indicator developped by Carayol et al. (2018). This indicator is also based on the frequency of the use of pairwise combinations of keywords, that are reported by the authors, in a given research field.[13] We define the variable *Novelty* as the maximum novelty score associated with a paper published in the three-years window preceding the

[13] See Appendix B for more details

date of the potential application. This variable may capture the ability of the researchers to define original research directions. Nevertheless, it does not reflect the specific novelty of the project, which is unobserved.

Various indicators have also been used in the literature to measure different concepts of interdisciplinarity (see Wagner et al. 2011, Rafols & Meyer 2009 for an overview). We consider that the degree of interdisciplinarity of a paper can be assessed by examining to what extent the paper relates to ideas developed in different disciplines. To built our indicator, we consider the disciplines assigned to the journals which publish the references cited by a paper. We calculate the Simpson diversity index[14] which describes the degree of diversity of disciplines[15] and is based on a combination of variety (number of categories) and balance (fair distribution among categories). We define the variable *Interdisciplinarity* as the mean interdisciplinarity score associated with articles published in the three years period preceding the date of the potential application. This variable may represent the skills the researcher has developed during his past experience when working on projects which combine multiple disciplines. However, it does not directly depict the degree of interdisciplinarity of the project.We hypothesize that these two variables may reflect the ability of the researcher to conceive non-conventional projects.

We also build a set of control variables which describe some factors that likely influence the submission of a proposal or/and the likelihood to receive a grant from the agency (Table 1.4). First, we use some individual characteristics of the researcher to control for different attributes, such as the age, the gender and the status. Considering that past scientific performances are particularly observed by the panel committees (Stephan et al. 2017), we define the variable *Citations* as the number of (three-years) citations[16] obtained by articles published in the three-year period preceding the date of the potential application, adjusted by the number of co-authors (fractional count). We use this variable to

[14]The Simpson diversity index is also known as 1-HHI, where HHI is the Herfindahl usual measure of concentration.

[15]Our main variable of interdisciplinarity uses the wide disciplines (aggregation of the subject categories of the Wos) associated with journals as categories. The ten categories are Life Sciences, Medicine, Ecology, Chemistry, Physics, Universe Sciences, Engineering, Mathematics, Social Sciences and Human Sciences. We also built a similar index which uses the 252 WoS subject categories instead of disciplines as level of aggregation, which is tested in the Robustness analysis Section V.4

[16]we also use alternative measures of the researchers' performances in the Robustness analysis Section V.4

measure the scientific impact of recent research. Given that future scientific performances are highly related to past performances, we also hypothesize that this variable may proxy the unobserved quality of the proposal, independently of its originality.

The quality of the laboratory to which the researcher is affiliated is likely to influence the probability to be awarded a grant. This may have a direct influence if the panel committee is sensitive to the reputation of the lab, or an indirect one as regards the research environment may stimulate the individual scientific performances. To represent the quality of the laboratory, we refer to the research units evaluation that the French Evaluation Agency (*HCERES*) makes every four years. We collect the grades associated with the overall quality of the researcher's laboratory,[17] which is assigned by the *HCERES* between 2007 and 2013 for most of the laboratories.[18]

We also consider that past accumulated experience in the submission of projects to the agency might give a learning advantage to researchers and influence the way they define their proposals. We create several variables that describe the participation in a program in the past: we use the number of projects submitted to the ANR in the past, a dummy which indicates if the researcher currently benefits from an ongoing ANR grant and a dummy that signals if the researcher submitted a project which was not awarded by the ANR in the previous year.

We also add in the equation which describes the decision to fund a project some variables which characterize the project. We measure the size of the team, its size squared and the duration of the project. We also add a dummy to control for the certification of a project by a competitiveness cluster.[19] We also control for the discipline of the researcher

and the year of application in each estimation. For the estimation of the Heckman probit selection model, a variable that influences the likelihood to submit a project but that is not correlated with the probability to obtain a grant is required. Thereby, to impose

[17]Our period under study covers 2005-2009 while we observe the laboratory of the researcher only once thereafter (around 2010). It is likely that some of the mobile researchers have had a different affiliation in the previous years, but we presume that these changes are small.

[18]We are not able to assign a grade for 21 research units since some evaluation reports only indicate comments about the lab and no score.

[19]It is likely that a project associated with a competitiveness cluster is more likely to be granted since the agency is presented as an important provider of financial support for these clusters. A cluster is defined in the scope of an industrial public policy program as a consortium of firms, research laboratories and educational institutions located in the same area and which collaborate on innovative projects.

Table 1.4 – Variables description used with the whole (or PI only) sample (for Tables 1.8 to 1.13)

Variables	Description
Dependent variables	
Project	=1 if the researcher submits a project to the agency
Grant	=1 if the applicant receives a grant
Independent variables	
Interdisciplinarity	average degree of interdisciplinarity by article
Novelty	maximum degree of novelty associated with an article
Individual characteristics	
Age	age
Age squared	age squared
Gender	=1 if male
Status	status (1='Assistant Prof.' (base), 2='Full Prof.', 3='Assistant Researcher', 4='Research Director')
Application experience	
Nb project	number of proposals previously submitted
Ongoing fund	=1 if the researcher benefits from an ongoing fund
Refusal	=1 if the researcher applied the preceding year and his project was not funded
Scientific production	
Citations	fractional count of the number of (three-years) citations received
Volume	fractional count of the number of articles published (robustness analysis only)
FI	highest Impact Factor associated with a journal in which the researcher has published (robustness analysis only)
Nb hit top10%	number of articles in the top 10% of the most cited articles in a discipline (robustness analysis only)
Lab characteristics	
lab size	number of tenured researchers and professors in the lab
Rank	global rank given to the lab by the French Evaluation Agency (1='A' (base), 2='A+', 3='B', 4='C' et 5='Missing')
Nb prev project lab	number of projects submitted in the past to the ANR by members of the lab
Current application	
Team size	size of the team
Team size squared	size squared of the team
Private part.	=1 if at least one partner of the project is associated with the private sector
Project duration	duration of the funding (in months)
Cluster	=1 if the project is associated with a competitiveness cluster
Control variables	
Discipline	research field from pooling of fine-grained sections(1='Life Sciences', 2='Medicine', 3='Chemistry', 4='Physics', 5='Universe Sciences', 6='Engineering', 7='Mathematics', 8='Information & Communication Sc. & Tech.', 9='Human and Social Sciences')
Year	date of application (1='2005' (base), 2='2006', 3='2007', 4='2008', 5='2009')
Program	ANR program (the eight programs are listed in Table 1.3)

exclusion restriction, we include an additional variable in the selection equation which describes the cumulative number of distinct projects submitted in the past by the members of the researcher's laboratory. There is no evidence that this variable can directly affect the ranking of submitted projects by the agency, but it can influence the likelihood to participate in several ways. First having colleagues that have already experienced an application to the ANR may have a positive or negative influence according to their opinion and also probably depending on whether they obtained a funding or not. This variable can also catch some incentives offered by the laboratory, such as some advice given to the members from the head or the employment of staff dedicated to help researchers to set up projects.

Team variables

The set of variable used in the model which refers to the teams are presented in Table 1.5. We use a sample which consists of all the teams having at least two members.[20] This time, our two independent variables which describe non-conventional research are formed from all the team members. We define *Novelty team* as the maximum of the novelty index associated with one member of the team in order to capture the ability of at least one investigator to set up an original project. We define *Interdisc. team* as the average interdisciplinary index in the team.

Following the work of Banal-Estañol et al. (2018), we also introduce some "inter-personal diversity" variables which capture the heterogeneity of the team members. It is argued that more diverse teams, defined as teams composed of researchers with different past experiences, diverse skills, different institutional affiliation or/and expertise from different research areas, may be more likely to turn into innovative research outcomes(Fleming 2001, Guimera et al. 2005, Jones et al. 2008, Post et al. 2009). Based on the Simpson variety index we previously used to measure interdisciplinary research, we calculate *Field diversity* to represent the diversity of disciplines in a team.[21] We refer to this variable to indicate multidisciplinary teams. Following the same approach, we construct three other covariates to characterize team diversity that may proxy the ability of the team to conceive original ideas. We define *Status diversity* measured as the diversity of status of the investigators, which is used to capture differences of ability according to researchers' career achievement or age. Another variable is named *lab diversity* and

[20] For instance, the non-directed programs dedicated to young researchers are not included in this sample since only one principal investigator is generally associated with a project

[21] From the known specialty of the researcher (his *section*), we are able to assign a discipline to each individual. We define nine categories: Life Sciences, Medicine, Chemistry, Physics, Universe Sciences, Engineering, Mathematics, Information & Communication Sciences and Technologies and Social and Human Sciences.

Table 1.5 – Variables description used with projects teams (Table 1.16)

Variables	Description
Dependent variables	
Grant	=1 if the the project is funded
Team and project characteristics	
Team size	number of scientific investigators in the team
Project duration	duration of the funding (in month)
PI male	=1 if the PI is a male
Age team	average age in the team
Cluster	=1 if the project is associated with a competitiveness cluster
Application experience	
Nb grant team	number of previously granted ANR-projects for the members of the team
Team scientific production	
Citations	average number of (three-years) citations (normalized by field) per investigator of the team
Interdisc. team	average degree of interdisciplinarity per investigator of the team
Novelty team	max degree of novelty in the team
Lab variables	
Top lab	=1 if at least one investigator is affiliated to a laboratory which was attributed the best grade A+
Heterogeneity of the team	
Field diversity	diversity in the disciplines of the investigators
Status diversity	diversity in the status of the investigators
Lab diversity	diversity in types of the laboratories of the investigators (under the responsibility of the CNRS, or another PRO, or the ministry)
Lab quality diversity	diversity in the scores attributed to the laboratories of the investigators
Control variables	
Year	date of application (1='2005' (base), 2='2006', 3='2007', 4='2008', 5='2009')
Program	ANR program (the eight programs are listed in Table 1.3)
Type of research	type of research of the project (1='basic research' (base), 2='experimental development', 3='industrial research')

defines the diversity of administrative supervisors of the laboratories involved in a team. Given that laboratories can be managed by multiple organisms, we simplify the present structure and define three mutually exclusive categories of type of laboratories: whether one of the supervisors is the CNRS; if not, whether one of the supervisors is a public research organism apart from the CNRS; and whether the laboratory is only under the supervision of a French Ministry. We also build the variable *Lab quality diversity* which represents the diversity of qualities of the laboratories involved in a team, based on the ranking performed by the French Evaluation Agency. Finally, we also introduce a dummy variable which represents whether the principal investigator is a male, to investigate the existence of gender bias in the grant process.

We finally define some variables, which are similar to the ones we use in the first model, to control for the characteristics of the investigators in the team. We use the average age of the members, the number of normalized citations the members received on average (fractional count), the total number of grants received in the past by the team members. We also define a dummy which represents whether (at least) one of the investigators is affiliated to a laboratory graded "A+" and a dummy to indicates whether the project is certified by a competitiveness cluster.

III.3 Descriptive statistics on the final sample

We present descriptive statistics on three distinct groups of our final sample: the non-applicants, the not-granted applicants and the granted applicants. The scientific investigators are included in the sample whatever their role in the team (PI or co-investigator). Table 1.6 allows us to observe many inter-group variations.

Individual characteristics

Researchers and faculty members that apply to the agency are slightly older on average than the non-applicants (around 44 years old versus 42.7 years old respectively). In our sample, scientists are mostly men (7/10 of non participants are men), this share being higher for the applicants and even more for the grantees (78% of granted investigators are men). Most of the non participants are faculty members rather than being full time dedicated to research (only 26% of researchers), while researchers are more numerous among the applicants. We have a fair distribution of researchers and faculty members in the funded group. While junior status (assistant professors and assistant researchers) are in majority in the non-applicant group (66%), scientists who apply and receive a grant (did not receive a grant) are more advanced in their career with 62% of full professors and researcher directors (respectively 55%). We see that the programs cover the main disciplines, but that their distribution is not well-balanced in our sample. We have a majority of Life Sciences scientists whatever the group (from 22% for the non-participant to 28% for the granted ones). Communication Science also has a high and steady share in the three groups (around 16%), followed by Chemistry (about 13%) and Medical Sciences (about 10%). The disciplines which account for less than 10% of the researchers are Mathematics, Universe Sciences and Human and Social Sciences.[22]

Scientific performances

We observe that applicants (to a larger extent granted ones) are characterized by better recent scientific performances than non applicants. They publish about 1 more article (in fractional counts) than the non-applicants on average (approximately 1.7 versus 0.91 articles). Their articles are also more cited, that is between 8.49 and 10.74 citations (in fractional counts) for the not-granted applicants and the grantees respectively, compared with only 3.99 citations for the non participants. They also tend to publish in higher Impact Factor (IF) journals; The best journal reached on average by granted applicants is characterized by an IF of 11, and only 5.8 on average for the non applicants. Granted researchers are also more likely to conduct research judged as excellent as regards the number of articles ranked in top 10% most cited articles in their specialty. They count 0.48 articles ranked in the top 10% on average, that is twice more than the non-participants. Finally, we see small differences between groups in terms of novelty and interdisciplinarity index. While applicants show higher scores than the non-participants for both novel and interdisciplinary research (about -0.33 vs. -0.46 and approximately 0.27 vs. 0.19 respectively), that is to say that they perform relatively more original and interdisciplinary research on average, funded investigators are not the ones that have the highest average scores. Nevertheless, the difference between not-granted applicants and grantees is slight (-0.30 vs. -0.35 and 0.28 vs. 0.26 respectively).

Laboratory characteristics

The research laboratories include about fifty researchers and faculty members on average. We see that applicants tend to be affiliated to relatively larger labs than non-applicants (between 53.3 and 55 members on average vs. 52.3 respectively). A larger share of the applicants is affiliated to a laboratory which is top ranked (“A+”) by the Evaluation agency and they are less numerous than non applicants in laboratories ranked “B”. Applicants tend to work more in laboratories where members are relatively more active in terms of

[22]The low number of SSH researchers can be explain by the removal of a large share of SSH disciplines after we collect the WoS publications

proposals submission to the ANR (but to a lesser extent if granted). It seems that the lab environment may influence the decision to apply or not.

Application experience

The applicants have already submitted one project on average to the agency in the past, that is more than the non-participants who applied 0.2 times on average (second part of Table 1.6, p. 36). They also already received more grants than the non applicants. This means that a share of the participants applies many times over the period, even if they already benefit from an ongoing funding.

Current application

The project team is composed of an average of 3.4 to 4 co-investigators and the project funding lasts three years on average. Concerning the type of partner in the team, we observe that only 0.02% of the applicant×project included at least one partner from the private sector in the proposal. While public/private partnership is not necessarily required in non-directed programs, this type of collaboration is often explicitly requested in directed programs calls. In addition, 21% of the grantees are associated with a competitiveness cluster, while it is only related to 2% of the not funded applicants.The programs

are characterized by different rates of applications. In our sample, we find that the most requested are non-directed programs (49.3% of the partner×projects). Among the directed-programs, *Biology and Heath* program attracts 20.3% of the investigators.

Composition of the team

Descriptive Statistics on the subsample constructed around teams and composed of 5,057 projects with at least two co-investigators are presented in Table 1.7. 33.2% of the submitted projects in this subsample is granted an award. We observe that whereas some variables are very close on average when comparing the not-funded teams to the granted ones (such as the average age, the total number of projects granted in the past), these two groups slightly differ in terms of recent past performances quality, interdisciplinary research and novelty, and quality of the laboratory. The teams of grantees receive on average slightly more citation (normalized) than the not-granted teams (0.57 vs. 0.38) and at least one of the investigator is affiliated to a top quality laboratory in 62% of the granted teams, compared to 55% of the not-granted. However scores associated with novel research and interdisciplinary research are both lower for funded teams (0.19 vs. 0.13 and 0.29 vs. 0.27 respectively). That means that they conduct less interdisciplinary research

Table 1.6 – Descriptive statistics on the groups of the non-participants, the not-granted applicants and the granted ones for the whole sample of researchers.

	Non-applicants		Not granted applicants		Granted applicant	
	mean	sd	mean	sd	mean	sd
Individual characteristics						
Age	42.71	9.72	44.11	8.08	43.93	8.06
Male	0.69	0.46	0.75	0.43	0.78	0.42
Status						
Assistant Prof.	0.49	0.50	0.23	0.42	0.17	0.38
Full Prof.	0.25	0.43	0.32	0.47	0.33	0.47
Assistant Resear.	0.17	0.38	0.22	0.41	0.22	0.41
Research Director	0.08	0.28	0.23	0.42	0.29	0.45
Discipline						
Life Sciences	0.22	0.41	0.27	0.45	0.28	0.45
Medicine	0.12	0.33	0.09	0.29	0.09	0.29
Chemistry	0.11	0.32	0.15	0.36	0.13	0.33
Physics	0.08	0.27	0.09	0.29	0.09	0.29
Universe Sciences	0.07	0.25	0.09	0.28	0.08	0.27
Engineering	0.09	0.29	0.09	0.28	0.09	0.29
Mathematics	0.09	0.28	0.03	0.17	0.04	0.20
Communication Sciences	0.16	0.36	0.15	0.36	0.17	0.38
SSH	0.07	0.25	0.03	0.16	0.03	0.16
Scientific production						
Citations	3.99	9.32	8.49	13.14	10.74	17.84
Volume	0.91	1.29	1.59	1.63	1.80	1.79
FI	5.75	8.41	9.88	10.69	10.95	12.40
Nb hit top 10%	0.21	0.71	0.40	0.96	0.48	1.06
Interdisciplinarity	0.19	0.21	0.28	0.20	0.26	0.20
Novelty	-0.46	1.24	-0.30	1.11	-0.35	1.13
Lab Characteristics						
Lab size	52.34	40.32	53.33	40.57	54.99	42.04
A+	0.48	0.50	0.51	0.50	0.47	0.50
A	0.37	0.48	0.40	0.49	0.47	0.50
B	0.13	0.33	0.07	0.26	0.05	0.23
C	0.01	0.08	0.00	0.04	0.00	0.04
Missing	0.01	0.11	0.01	0.12	0.01	0.11
Nb prev project lab	7.72	11.32	12.15	13.30	9.31	12.02
Observations	118,431		14,453		6,610	

Table 1.6 Continued

	Non-applicants		Not granted applicants		Granted applicant	
	mean	sd	mean	sd	mean	sd
Application experience						
nb projects	0.19	0.60	1.06	1.50	0.89	1.43
nb grants	0.08	0.32	0.29	0.63	0.29	0.65
Ongoing grant	0.03	0.18	0.09	0.28	0.07	0.26
Refusal	0.04	0.19	0.32	0.47	0.27	0.44
Current application						
Team size	.	.	3.39	1.93	3.95	2.33
Project duration	.	.	38.06	6.03	37.22	6.14
Private part.			0.22	0.41	0.29	0.45
Cluster	.	.	0.02	0.16	0.21	0.41
Control variables						
Program						
Biol.&Health	.	.	0.20	0.40	0.21	0.40
Energy	.	.	0.02	0.12	0.08	0.27
Environment	.	.	0.07	0.25	0.10	0.31
Security	.	.	0.03	0.16	0.02	0.15
Materials	.	.	0.08	0.27	0.08	0.27
Non-directed	.	.	0.54	0.50	0.40	0.49
SSH	.	.	0.01	0.12	0.01	0.10
Communication	.	.	0.06	0.23	0.10	0.30
Year						
2005	0.21	0.41	0.05	0.21	0.20	0.40
2006	0.19	0.40	0.23	0.42	0.22	0.42
2007	0.20	0.40	0.22	0.42	0.21	0.41
2008	0.20	0.40	0.17	0.38	0.18	0.39
2009	0.20	0.40	0.33	0.47	0.18	0.38
Observations	118,431		14,453		6,610	

on average and that the investigator which performs the most novel research in the team has a novelty score lower than not-funded teams on average. The team of grantees is composed of a total of 4.3 co-investigators on average, while the non funded has 3.9 partners and granted teams are much more often associated with a competitiveness cluster (21% vs. 3% respectively). The distribution of the type of research is not balanced in the sample. Basic research is associated to 83.6% of the projects, 14% of the projects do industrial research and the remaining 2.4% perform experimental development. The difference between not-granted and granted in the type of research is very low.

When we turn to the variables which describe the heterogeneity of the team, we observe that the diversity index are on average slightly lower for granted teams, apart from the variety of status and the diversity of laboratories' quality which are equal between the two subgroups. This means that the configuration of the not-funded teams are on average slightly more multidisciplinary than granted ones (0.31 vs. 0.28 respectively), and are also more diverse in terms of the combination of different type of laboratories (0.15 vs. 0.13 respectively).

IV Methodology

We model the application and success processes using two equations, each of them is a binary choice model. The selection equation describes the decision of a researcher to submit a project to the funding agency while the structural equation describes the decision of the agency to grant a project. It is likely that the evaluation a researcher makes on its own project when he takes the decision to apply is correlated with the evaluation of the agency, which implies a correlation between the two binary outcomes. In this case, some unobserved covariates can probably affect both decisions and the implementation of two independent probit models would biased the estimates (Heckman 1976, 1979). We consider this problem and estimate by maximum likelihood a Heckman probit selection model which is a join model taking into account selection bias, and allows the errors terms of the two equations to be correlated. In this way, u_i and v_i, the error terms of the first and second equations respectively, follow a bivariate normal distribution. We first present these two models separately, and describe the joint model in the last subsection.

Table 1.7 – Descriptive statistics on the groups of the not-granted applicants and the granted ones with the sample of teams.

	Not granted		Granted	
	mean	sd	mean	sd
Team variables				
Age team	45.02	5.64	44.34	5.90
PI male	0.75	0.43	0.78	0.42
Citations team	0.38	0.94	0.57	1.10
Interdisc. team	0.29	0.15	0.27	0.15
Novelty team	0.19	0.54	0.13	0.58
Top lab	0.55	0.50	0.62	0.49
Nb grants team	0.77	1.13	0.74	1.19
Team size	3.94	1.71	4.35	2.13
Project duration	38.24	5.82	37.53	5.92
Private part.	0.23	0.42	0.29	0.46
Cluster	0.03	0.16	0.21	0.41
*Type of research**				
experimental development	0.02	0.15	0.03	0.16
basic research	0.86	0.35	0.79	0.41
industrial research	0.12	0.32	0.18	0.39
Members diversity				
Field diversity	0.31	0.25	0.28	0.25
Lab diversity	0.15	0.23	0.13	0.22
Lab quality diversity	0.26	0.25	0.26	0.24
Status diversity	0.37	0.23	0.36	0.24
Control variables				
Program				
Biol.&Health	0.22	0.42	0.20	0.40
Energy	0.02	0.13	0.08	0.28
Environment	0.08	0.27	0.11	0.31
Security	0.03	0.17	0.03	0.16
Materials	0.09	0.29	0.09	0.28
Non-directed	0.49	0.50	0.38	0.49
SSH	0.01	0.09	0.00	0.05
Communication	0.07	0.25	0.11	0.32
Year				
2005	0.05	0.22	0.18	0.39
2006	0.24	0.43	0.23	0.42
2007	0.23	0.42	0.22	0.41
2008	0.16	0.37	0.19	0.39
2009	0.32	0.47	0.18	0.38
Observations	3,378		1,679	

*Notes: * The type of research of the project is only known for 4,670 projects.*

IV.1 The decision of the researcher to submit a project to the Agency

The decision of the researcher is a binary response, set to 1 if he decides to apply or 0 otherwise. We model his decision with the following equation:

$$A_{it}^{*} = \alpha_1 X_{i,t}^{1} + \alpha_2 X_{j,t}^{2} + \mu_{i,t} \tag{1.1}$$

where A_{it}^{*} is the unobserved latent variable, $X_{i,t}^{1}$ represents a set of individual characteristics and $X_{j,t}^{2}$ a set of information related to the laboratory of the researcher. These variables are described in Table 1.4. $\mu_{i,t}$ is the error term.
The latent variable represents how a researcher values his research project. We cannot observe directly the true value of a project, but we hypothesize that it depends on some characteristics of the researchers and is especially correlated with his past scientific production.

The binary response of the model is obtained as follow:

$$A_{it} = \begin{cases} 1 & if\ A_{it}^{*} > 0 \text{ if the researcher decides to submit a project to the agency} \\ 0 & if\ A_{it}^{*} \leq 0 \text{ otherwise} \end{cases}$$

If $A_{it}^{*} > 0$, the researcher considers that the value of his project his sufficiently high so that he should have a good chance to be awarded a grant, and decides only in this case to apply.

In the case of the estimation of a standard probit model, we can estimate the probability that a researcher apply by :

$$Pr\left[A_{it} = 1\right] = Pr\left[A_{it}^{*} > 0\right] = Pr\left[\alpha_1 X_{i,t}^{1} + \alpha_2 X_{j,t}^{2} + \mu_{i,t} > 0\right] = \Phi\left(\alpha_1 X_{i,t}^{1} + \alpha_2 X_{j,t}^{2}\right)$$

where $\Phi(.)$ is the cumulative univariate distribution function for the standard Normal distribution.

IV.2 The decision of the agency to fund a project

The decision of the agency is also a binary response, set to 1 if he decides to select and grant a project or 0 otherwise. We model his decision with the following equation:

$$G_{it}^{*} = \beta_1 X_{i,t}^{1} + \beta_2 X_{j,t}^{2} + \beta_3 X_{p}^{3} + \beta_4 X_{k,t}^{4} + \upsilon_{i,t} \tag{1.2}$$

where G_{it}^{*} is the unobserved latent variable, $X_{i,t}^{1}$ and $X_{j,t}^{2}$ are respectively the same sets of individual characteristics and laboratory information, as the ones included in equation 1.1. X_{p}^{3} represents the type of program and some information related to the project are included in $X_{k,t}^{4}$. $\upsilon_{i,t}$ is the error term.

The latent variable represents the assessment of the project quality by the agency. Given that we do not know the grades assigned to the projects by the panel committees, we cannot directly observe this value. We consider that the result of the evaluation is correlated with some characteristics of the applicants and of the project.

The binary response of the model is obtained as follow:

$$G_{it} = \begin{cases} 1 & if\ G_{it}^{*} > 0 \text{ if the agency decides to grant a project} \\ 0 & if\ G_{it}^{*} \leq 0 \text{ otherwise} \end{cases}$$

If $G_{it} > 0$, the agency judges that the value of the project is high enough and decides only in this case to award a grant to the applicants.

In the case of the estimation of a standard probit model, we can estimated the probability that the agency awards a grant by :

$$\begin{aligned} Pr\left[G_{it}=1\right] = Pr\left[G_{it}^{*}>0\right] &= Pr\left[\beta_1 X_{i,t}^{1} + \beta_2 X_{j,t}^{2} + \beta_3 X_{p}^{3} + \beta_4 X_{k,t}^{4} + \upsilon_{i,t} > 0\right] \\ &= \Phi\left(\beta_1 X_{i,t}^{1} + \beta_2 X_{j,t}^{2} + \beta_3 X_{p}^{3} + \beta_4 X_{k,t}^{4}\right) \end{aligned}$$

where $\Phi\left(.\right)$ is the cumulative univariate distribution function for the standard Normal distribution.

IV.3 The connected model

We estimate two probit models with sample selection, in which the errors terms u_i and υ_i follow a bivariate normal distribution, with null mean and variance-covariance matrix equals to $\left(\begin{smallmatrix} 1 & \rho \\ \rho & 1 \end{smallmatrix}\right)$ to allow the distribution of the errors terms to be correlated. According to the responses in each equation, we can consider the three different cases $\left(A_{it}, G_{it}\right) \epsilon \left\{\left(0,0\right), \left(1,0\right), \left(1,1\right)\right\}$. The associates probabilities can be written as:

$$\begin{aligned} &Pr\left[A_{it}=0, G_{it}=0\right] = \Phi\left(1-\left(z_1\right)\right) \\ &Pr\left[A_{it}=1, G_{it}=0\right] = \Phi\left(z_1\right) - \Phi_1\left(z_1, z_2, \rho\right) \end{aligned}$$

$$Pr\left[A_{it}=1,G_{it}=1\right]=\Phi_1\left(z_1,z_2,\rho\right)$$

where $z_1 = \alpha_1 X^1_{i,t} + \alpha_2 X^2_{j,t}$ and $z_2 = \beta_1 X^1_{i,t} + \beta_2 X^2_{j,t} + \beta_3 X^3_p + \beta_4 X^4_{k,t}$. $\Phi\left(.\right)$ is the cumulative univariate distribution function for the standard Normal distribution and $\Phi_1\left(.\right)$ is the cumulative bivariate Normal distribution.

V Estimations and results

We estimate how novel research and interdisciplinary research performed by a researcher in the past influence his decision to submit a project to the ANR, and affect the probability of the applicants to be granted with a Heckman probit selection model. This model allows us to correct for the selection bias that could have biased our results if we would have not specified that the sample of applicants is drawn from a larger population. We first present the main results estimated from the sample which includes all the applicants, and then those when we limit the sample to the PI only. In each specification, the estimation of the standard errors allows correlation between repeated observations of each non-applicants and between members of a team for the applicants. We report at the bottom of the results tables the rho coefficient which measures the degree of correlation between the selection and structural equations and add the output of a Wald test of independent equations. In a few cases, we cannot reject at the 5% level the null hypothesis of absence of correlation between the two equations, which means that we could have used two independent probit models instead. For all of the estimations, we report the mean marginal effects associated with the covariates, besides the estimated coefficients of the probit model are presented in Appendix D.

V.1 Main results

The mean marginal effects associated with the estimation of the influence of novelty and interdisciplinary research on the probability to apply are reported in Table 1.8 and in Table 1.9 when we estimate the conditional probability to receive a grant. We first introduce the control variables which describe the researchers' characteristics in the two first specification (columns 1 and 2) and add characteristics of the laboratory[23] in the two following specifications (columns 3 and 4). The full specification is finally reported in column 5 given that novelty and interdisciplinarity are weakly correlated.

First, we find that the control variables display expected relationship with the de-

[23]While the laboratory environment may have some influence on the researcher's behavior, we observe the researcher's affiliation around 2010.

pendent variables. The quality of the recently performed research influences positively and significantly both the probability to apply and the conditional probability to get a grant. A Researcher who gets 20 more citations for some articles co-authored with one person increases the likelihood to apply and to get funded by 2%, all other factors being equal. Whereas men tend to apply more but at a very low extent (+1.3%), gender does not affect significantly the funding decision. We also see that researchers more advanced in their career have a higher probability to apply but it does not affect significantly the grant process. Professors are approximately 9.4% more likely to submit a project than associate professors. Researchers employed by a public research organism are also more likely to participate than faculty members, with 7.7% (respectively 15.2%) more chances to apply for associate Researchers (Research directors) compared to assistant professors. While the age of the researcher is associated to a negative mean marginal effect for both decisions to apply and to get a grant, its effect rather follows an inverted-U shape in the selection equation (Table 1.22 in Appendix D) and a U shape in the structural equation (Table 1.23).

The lab environment also influences both the researcher and agency decisions. Researchers affiliated to laboratories which were not well evaluated by the Evaluation Agency are less likely to apply than researchers working in laboratories graded “A” or “A+”, *ceteris paribus*. To be affiliated to a laboratory ranked “B” (“C”) decreases the likelihood to apply by 3.2% (respectively 6.2%). This difference is even clearer when we refer to the decision to fund a project. This time, only top laboratories’ applicants (“A+”) are favored, with 3.6% more chances to get a grant than applicants from a laboratories ranked “A” *ceteris paribus*. However the size of the lab has no significant influence on the outcomes.

We also find that the participation in ANR programs affects future applications to the agency. The submission of one more project in the past increases by 4.2% the likelihood to apply again, but does not affect significantly the chances to get a grant. Researchers who still benefit from an ongoing ANR funding received in the past have 5.5% less chances to submit a new proposal while those who apply and have not been funded are 16.4% more likely to participate in the subsequent year.

We also add some characteristics of the project in the structural equation. We find that the size of the team has a positive and significant effect on the likelihood to get a grant. The effect in fact follows an inverted-U shape (Table 1.23). This variable is often related in the literature to a proxy of the heterogeneity of a team. We investigate in subsection V.3 the diversity of the team assembly to test this idea in detail. The type of partners involved in the team has a significant but low weight on the funding decision. Having at least one partner from the private sector in the team decreases the chances to get a grant of 2.7%, *ceteris paribus*. The effect of the project duration is non different from zero. We also find that the agency tends to highly favor applicants whose project is associated with a competitiveness cluster, with an increase of 44% of the likelihood to be awarded a grant.

When we turn to the estimated influence of our independent variables on the different outcomes, we first find a positive and significant relationship between recently performed interdisciplinary research and the probability to participate in a call for proposal, *ceteris paribus*. It is likely that the design of some of the ANR programs and the priority areas defined by the agency attract more the researchers who already experienced relatively higher degree of interdisciplinary research. Then, we also find that research novelty affects positively and significantly the likelihood to submit a project to the agency. Once again this seems to go the right way in suggesting that the claims of the programs have well been understood by the scientific community and that researchers, whose recent work is relatively more novel, are more likely to apply to the agency. As previously seen, researchers with better recent past performances, those more advanced in their career or affiliated to a laboratory characterized by a high quality of the scientific environment may value more their work than less recognized scientists and thus are more likely to apply. Then, all the aforementioned factors being equal, researchers who selected more original research direction or carried out interdisciplinary research recently seem to assess even more favorably their work than scientists who perform more conventional research. In fact, they are more likely to apply since they may believe that their works is high quality and deserve to be awarded a grant.

However, when we turn to the conditional probability of the applicants to be awarded a grant, we find that researchers associated with novel or interdisciplinarity research are less likely to receive a grant, holding research quality and other factors constant. These results seem not to match what the agency is looking for and support our hypothesis that researchers who performed non-conventional research recently tend to be disadvantaged during the evaluation process. These results are also valid when we introduce the full set of control variables in the model (from columns 3), and hence control for the quality of the laboratory and some characteristics of the project. It seems that the agency, which has less information about the project than the applicants, does not judge the proposal in the same way as applicants do with their private information. These projects may be more difficult to assess, and the uncertainty of the outcome can drive the reviewers not to go ahead with the project. Researchers may also be penalized if past publications of applicants are influential on the decision of the committee and if reviewers consider that it is less risky to fund researchers who are used to perform conventional research.

Our results still hold when we limit our sample of applicants to the Principal investigators only (Table 1.10 and Table 1.11). Novel research and interdisciplinary research are still associated with a lower conditional probability to get a grant, whereas researchers who used to conduct these type of non-conventional research are more likely to apply in the first step.

V.2 Differentiation according to the type of programs

The difference in the design of programs allow us in a second time to refine the analysis and to differentiate the results according to the type of programs offered by the agency. We estimate the same models as in the previous subsection for both types of directed and non-directed programs separately. Our samples of applicants are fairly well balanced between the two types of programs. We count 7,322 projects submitted to the non-directed programs (10,375 partner×project) and 6,810 projects submitted to the directed programs (10,685 partner×project).

We find that directed funding programs attract relatively more male researchers, with better career achievement, whose publications impact are higher and that are affiliated to good research laboratories (Table 1.12). Full professors (research director) are 5.6% (8.2% respectively) more likely to apply to a directed program than assistant professors. Researchers who get ten more citations for sole author papers have 1% more chances to apply, which is a rather low effect. Moreover, the researchers affiliated to labs graded "A" are more likely to apply than others, even more than members of top labs graded "A+", all other factors being equal. The effects associated with variables reflecting the application experience go in the same directions as the ones obtained in the preceding model. Finally, novel and interdisciplinary research have a positive and significant effect on the probability to apply to the directed programs,*ceteris paribus*. However, holding research quality and other factors constant, we observe again a negative and significant effect of non-conventional research on the conditional probability to be granted (Table 1.13). Researchers whose parts of their research are based on more original ideas or those who perform higher degree of interdisciplinary research are less likely to have their project granted. Unlike the previous model, the project duration has a low but negative and significant effect on the likelihood to receive a grant. Moreover, researchers affiliated to top-ranked laboratories are 4% more likely to get a grant, and the effect associated with the impact of research is slightly higher than in the selection model (+2% if ten more cites obtained).

When we turn to the investigation of non-directed programs (Table 1.14), control variables affect the probability to participate in the same way as the preceding model, apart from the fact that researchers affiliated to top-ranked laboratories are slightly more likely to apply than others (+0.7% compared to labs graded “A”). Although interdisciplinarity still has a positive and significant effect, we find that novel research does not influence significantly the likelihood to participate in non-directed programs. Moreover, the research originality is also always non-significant at the 5% level in the structural equation whereas a higher degree of interdisciplinarity is still associated to a negative and significant effect on the grant process (Table 1.15). We also observe that having a partner from the private sector does not have any significant influence on the funding decision for both type of programs.

These results contrast with the main ones in several dimensions. Whereas it is stated in non-directed programs calls that novelty and originality of the projects are particularly expected, this criterion does not significantly influence researchers in their decision to apply. Concerning the decision of the agency to grant a projects, there is still no evidence that having carried out original projects is influential during the evaluation projects. One possible explanation for the observed bias against interdisciplinary research could be that projects are evaluated by disciplinary review panels, who might be less able to assess the potential impact of interdisciplinary research project. However, it should not be the only reason which explains this result since projects are evaluated by multidisciplinary committees in the directed programs while a negative effect is also found. High priority areas defined by the agency in the directed programs seem therefore more adapted to attract researchers who have performed highly innovative research, compared to the non-directed programs, although these researchers are then penalized during the grant process. On the other hand, the non-directed programs seem to attract relatively more researchers affiliated to renown laboratories, even though this effect is very low.

V.3 Heterogeneity of the team

We then test if our main results are consistent with the ones obtained when we consider variables measured at the level of the team instead of individual variables. On top of that we also investigate how several measures of team diversity are associated with the decision of the agency to eventually fund a project. In this case, the results are obtained from the estimation of a standard probit model (probit coefficients are reported in Table 1.30).

We investigate to what extent the degree of research conformism of the team members and the specific team composition are associated with the grant process using our subsample of 5.057 projects. We first test our independent variable along with our regular control variables and then introduce the type of research of the proposal as control variable. This variable is included later in the model since we only know this information for a subset of projects (that is for 4,670 projects). We first find that the relationship between the control variables and the decision to fund a project is consistent with our preceding results (Table 1.16). The mean marginal effect associated to the average age of the members is negative and significant, the effect having an U-shape according to the mean age (1.30). The number of citations (normalized by field) obtained on average in the team, that we use as a proxy for the project quality, is positively and significantly associated with the funding decision. Moreover, teams composed of at least one scientific investigator affiliated to a top-ranked laboratory ("A+") increase the probability to get a grant of 6%, all other factors being equal. The total number of ANR funding obtained in the preceding years by the members of the team has a positive relationship with the dependent variable, that is one additional funding obtained increases the probability of the project to be successful of 1.7%. However, we find no significant influence either for a public/private partnership nor for the size of the team. Moreover, the same applies for the duration of the project and the gender of the PI that both display not significant coefficients. Finally, the distribution of the type of research is not balanced since basic research is associated to 83.6% of the projects in our sample (14% in industrial research and 2.4% in experimental development), and doesn't not have any significant direct effect on the grant process.

Our independent variables give some results that are consistent with the ones we previously obtained. Both the interdisciplinarity and the novelty index measured at the level of the team are negatively and significantly associated with the likelihood of success, *ceteris paribus* (column 1-3 in Table 1.16). These results give confirm the idea that projects whose researchers' team have performed relatively more non-conventional research are negatively biased in the grant process. However, the result associated to interdisciplinary research is no longer significant at the 5% level when we introduce the type of research of the project as control variable (column 4 in Table 1.16) while the result for novelty still hold.We also find that most of our indicators used to capture the heterogeneity of

the team are significant and negatively associated with the probability of a project to be successful, whatever the specification, and *ceteris paribus*. Thus, the more the teams are diverse, the less their likelihood to receive a grant. Only the diversity of the team in terms of the status of the investigator has a negative but not significant effect on success. In addition to the negative effect of the conduction of interdisciplinary research by investigators taken individually, we find that multiple disciplines teams, hence multidisciplinary ANR proposals, are less likely to be granted when the number of disciplines involved on a project increases (*Field diversity* variable). The diversity of the type of labs, defined by the type of institution in charge of the administrative supervision of the lab, has also a negative and significant effect on the probability of success. We also find a negative effect for the variety of lab quality. These two variables may indicate more diverse teams in terms of knowledge and expertise since these teams combine various research institutions and different quality of laboratory environment. These results give some evidences that in addition to the negative effect of non-conventional research conducted by the members separately, teams that are more diverse in terms of knowledge and expertise of the members are penalized during the grant process.

V.4 Robustness analysis

We modify the specification of our model in order to verify we did not miss any influential control variables (Tables presented in Appendix C).

Since the results may be sensitive to the choice of the indicator, we first defined other measures of interdisciplinarity. We built a new variable based on the Shannon-Wiener index. As compared to our main interdisciplinarity index used, this indicator of diversity has the distinctive feature to put more weight on less frequently used disciplines. We also construct the two indexes using specialties (WoS subject categories) instead of disciplines as level of aggregation. We find that the main results holds with these alternative indicators and that researchers whose work is more interdisciplinary are still more likely to submit a proposal to the agency (Table 1.17) whereas interdisciplinary research is still negatively and significantly related with the conditional probability of success (Table 1.18). We observe that the effects associated with the indexes are slightly lower when constructing the indexes using specialties (Wos subject categories) instead of disciplines, which implies a larger number of categories.

We present another model in which we substitute the control variable *Citations* with some other measures of scientific performances: a fractional count of the number of publication (*volume*), the ability to publish paper in well established journals (*IF*, the maximum journal impact factor) and excellence of the research (*nb hit 10%*, that is the number of articles among the 10% most cited). Some of these measures are highly correlated, so we introduce them separately, but they may also capture different dimensions of research performances that we need to control for. The results obtained with these new specifications do not alter our main results on novelty and interdisciplinarity indexes since the still have a positive relationship with the application decision and a negative one with the grant decision (Table 1.19 and Table 1.20). All the research performances variables are positively and significantly associated with the decision to apply, although the effect is relatively low. For instance, two more (sole authored) publications increase the likelihood to apply by 2.2% (to 3.2%), whereas one more top cited publication is associated to an increase of 0.3%. Nevertheless, only the ability to reach high impact factor revue influences significantly at the 5% level the grant decision (*volume* is also significant, but only in half of the specifications). In the last two specifications, we add the type of research associated with project as additional control variable to our main specification model (the full set of individual covariates and the characteristics of the laboratory). Our sample is then reduced to 19,471 partner×project, and we find that our results related to novel research and interdisciplinarity are still valid.

Finally, programs components are characterized by different levels of competition for the selection of proposals. Some directed programs components are characterized by a relatively low number of submitted proposals, that permits a larger share of the applications to be funded. We verify that our results are not influenced by some programs characterized by a high success rate, which could imply that less relevant projects are also funded, we discard the applications to specific calls with a success rate above 80% in a first specification and above 60%[24] in a second specification. We also add a categorical variable to control for each distinct call and each year. We can not observe significant changes in the results compared to our first model, novelty and interdisciplinarity are still negatively and significantly associated with the funding decision (Tables 23 and 24). Applications in programs with small number of submitted applications and a high success rate, which could have been less original, do not seem to alter our results.

VI Conclusion

This chapter contributes to the literature on the allocation of competitive funds to researchers by funding agencies. We especially focus on whether researchers who have addressed more novel approaches or have performed more interdisciplinary research self-select and whether they are then selected for funding. We also investigate how the composition of the team, measures in terms of diversity of disciplines and cognitive diversity of the members, influence the probability of an application to be successful. Our data set is composed of a large sample of French researchers and professors, along with some information about their successful and unsuccessful participations to a program launched by the French funding agency. We model a heckman probit selection model to assess which factors are significantly associated with the decision of the agency to fund a project, while controlling in a first step for the factors that encourage the researchers to apply.

We first find some evidence of a self-selection process, in which the best researchers in terms of past research performance and renown of their laboratory are more likely

[24]These high success rates are also artificially obtained because our final sample only include 60% of the submitted projects over the period 2005-2009.

to apply. Our results also suggest that researchers who selected more original research directions or carried out interdisciplinary research in the recent past are more likely to participate in a funding program, the research quality and all other factor being equal. Everything works as if these researchers who perform non conventional research tend to value more favorably their work than scientists who are more conformist, and they may think that their qualitative work has good chance to be granted. We also find that projects whose investigators performed non-conventional research during the preceding years are less likely to be granted, while controlling for the quality of the research and characteristics of the investigators. We also find that some indicators which describe the heterogeneity of the investigators in a team are negatively and significantly associated with the conditional likelihood of success. These indicators reflect the variety of disciplines involved in a project, that is to say the interdisciplinarity of a project, and some cognitive diversity of the teams members. These results run counter to one of the main missions of the agency, which consists in funding innovative research, and suggest that the selection of projects can be biased by reviewers' risk-averse behaviors. In a context of information asymmetry, evaluators may not value the potential impact of some projects in the same way as applicants do, since they may not know the private information that applicants have. In this case, projects based on more conventional ideas may be favored by the committees since reviewers predict a greater feasibility along with less uncertainty over their results. This argument would support the hypothesis that researchers who performed non conventional research in the past are creative and are then more likely to conceive original projects in the future, although it is not proved. An other explanation could be that reviewers identify the research originality of applicants when they look their list of past publications and CV up, and then could penalize the researcher more likely to carry out risky projects.

As presented in the literature, information reflecting that the researcher is promising, such as the number of citations received or the quality of his research laboratory, are also influential in the grant process. It is also very likely that this information is correlated with the quality of the submitted project, independently of its originality.

We also perform the analysis having divided the sample in two parts, according to the type of programs. The agency preliminary defines the subject area of the projects for the directed programs while researchers can freely define their research issue before the submission of the project to non-directed programs. The former are also often more suitable for applied or industrial research, compared to the latter which are more aimed at funding basic or emerging research. We find that the likelihood of success is still negatively biased for researchers who perform original and interdisciplinary research in the directed programs. However, we also observe that the design of these programs is well defined as they attract more original and interdisciplinary researchers. When we consider non-directed programs, we find that interdisciplinarity is negatively and significantly associated with the likelihood of success. This can be explained by the fact that review panels are composed of disciplinary researchers who may be less able to assess the quality of interdisciplinary research project. However this does not hold for directed programs because projects are evaluated by multidisciplinary committees. Finally we find for this type of programs that the association between research novelty and the funding decision is not significantly different from zero at the 5% level. However researchers who apply to non-directed programs do not perform more novel research than those who do not apply to this type of program. Thereby directed programs, which target priority subject areas, seem more adapted to attract researchers who design highly innovative projects than non-directed programs. However, we can not know which type of program finally fund the most novel research. On one hand, researchers who participate in directed programs may be more creative and may be more likely to conceive original projects, but the most original ones are eventually not funded by the committees. On the other hand, the applicants to the non-directed programs are researchers who are used to conceive more conventional projects and we do not find clear evidence that the reviewers penalize novel research.

The major limitation of our analysis is that we do not observe the quality, originality and interdisciplinarity of the proposals directly. We presume that future research performances are correlated with previous ones, so that the past research performances may give some information about the ability of the researchers to conceive a good project. It is also conceivable that researchers who have previously conducted non-conventional research may redo it, given that they are not risk-averse and may be more inclined not to restrain their creativity. Our measure of interdisciplinarity of the project, which uses the disciplines of the applicants, may be relatively correct. Another limitation is that we study several funding programs all together. While most of the programs share the explicit requirement to address original questions and the possibility to make up multidisciplinary teams, each call launched by the agency has its own design, along with potentially different eligibility rules and claims to conceive the project. Some programs for instance require to involve both public and private partners in the project team, whereas others are more appropriate for basic research. Some programs also specifically target young researchers. We thus may have failed to control for some important factors which influence the selection of projects by the agency. Given that the size of the subsamples according to the specific type of programs do not allow us to investigate accurately the grant process program by program, we only perform the analysis on two different type of programs separately.

Given that we found some evidence that the design of the funding program may influence which type of researchers will participate in, even if some eligibility criteria are not stated, it may have several applications. First, agencies should carefully decide the design of their programs since it may have some implications on the researchers who participate in and thus on the results. Then, when studying the impact of funding on ex-post performance of the grantees, it is important to clearly understand to what extent the applicants differ from the reference population, in order to avoid inaccurate generalized interpretations of the efficiency of the grant allocation.

Finally, this study only refers to the first five years of activity of the ANR, so we cannot generalize our results in asserting that they are still valid. Some changes seem to have occurred until today, for instance in the definition of funding programs and in the evaluation process. One direction for future research could be to investigate the drivers of a successful application on a longer period, in order to test if the agency adjusted his allocation mechanism over time.

Table 1.8 – Factors that influence the probability to submit a project for the whole sample (Mean marginal effects reported)

	(1)	(2)	(3)	(4)	(5)
Novelty		0.003***		0.003***	0.005***
		(0.005)		(0.005)	(0.005)
Interdisciplinarity	0.093***		0.092***		0.096***
	(0.028)		(0.028)		(0.029)
Individual variables					
Age	-0.001***	-0.001***	-0.001***	-0.001***	-0.001***
	(0.006)	(0.005)	(0.006)	(0.006)	(0.006)
Male	0.013***	0.015***	0.013***	0.015***	0.013***
	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)
Status					
Professor	0.094***	0.097***	0.094***	0.097***	0.094***
	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)
Associate Researcher	0.080***	0.081***	0.077***	0.078***	0.077***
	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)
Researcher	0.156***	0.157***	0.152***	0.154***	0.152***
	(0.022)	(0.022)	(0.022)	(0.022)	(0.022)
Year					
2006	0.089***	0.090***	0.089***	0.091***	0.089***
	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)
2007	0.040***	0.041***	0.042***	0.043***	0.042***
	(0.021)	(0.021)	(0.021)	(0.021)	(0.021)
2008	0.003	0.006	0.007	0.008*	0.007
	(0.023)	(0.023)	(0.023)	(0.023)	(0.023)
2009	0.044***	0.047***	0.048***	0.050***	0.048***
	(0.022)	(0.022)	(0.023)	(0.023)	(0.023)
Citations	0.002***	0.002***	0.002***	0.002***	0.001***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Application experience					
nb projects	0.042***	0.043***	0.042***	0.043***	0.041***
	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)
Ongoing grant	-0.054***	-0.055***	-0.055***	-0.056***	-0.055***
	(0.024)	(0.024)	(0.024)	(0.024)	(0.024)
Refusal	0.164***	0.167***	0.164***	0.167***	0.164***
	(0.019)	(0.019)	(0.019)	(0.019)	(0.019)
Lab variables					
Lab size			0.000	-0.000	0.000
			(0.000)	(0.000)	(0.000)
Nb prev project lab	0.001***	0.001***	0.000**	0.000***	0.000**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Rank					
A+			0.003	0.002	0.003
			(0.013)	(0.013)	(0.013)
B			-0.032***	-0.035***	-0.032***
			(0.020)	(0.020)	(0.020)
C			-0.062***	-0.065***	-0.062***
			(0.100)	(0.099)	(0.100)
Missing			0.011	0.012	0.012
			(0.051)	(0.051)	(0.051)
Scientific field	YES	YES	YES	YES	YES
Observations	139,494	139,494	139,494	139,494	139,494

Notes: Robust standard-errors in parentheses, clustered at the individual or project level. The control variables Scientific field is included in the model but the mean marginal effects are not reported.

Table 1.9 – Factors that influence the probability to receive a grant for the whole sample (Mean marginal effects reported)

	(1)	(2)	(3)	(4)	(5)
Novelty		-0.008***		-0.006***	-0.007***
		(0.009)		(0.009)	(0.009)
Interdisciplinarity	-0.068***		-0.070***		-0.073***
	(0.067)		(0.068)		(0.070)
Individual variables					
Age	-0.003***	-0.003***	-0.003***	-0.003***	-0.003***
	(0.017)	(0.018)	(0.017)	(0.018)	(0.017)
Male	0.014	0.013	0.010	0.009	0.010
	(0.024)	(0.024)	(0.025)	(0.026)	(0.025)
Status					
Professor	0.054	0.053	0.048	0.049	0.049
	(0.069)	(0.071)	(0.066)	(0.072)	(0.066)
Associate Researcher	0.058	0.058	0.056	0.057*	0.057
	(0.063)	(0.063)	(0.058)	(0.063)	(0.059)
Researcher	0.103	0.103	0.094	0.095	0.094
	(0.106)	(0.107)	(0.098)	(0.107)	(0.099)
Year					
2006	-0.298***	-0.298***	-0.308***	-0.308***	-0.307***
	(0.049)	(0.053)	(0.053)	(0.057)	(0.053)
2007	-0.306***	-0.306***	-0.386***	-0.387***	-0.385***
	(0.051)	(0.050)	(0.060)	(0.059)	(0.060)
2008	-0.292***	-0.292***	-0.355***	-0.356***	-0.354***
	(0.059)	(0.055)	(0.067)	(0.064)	(0.067)
2009	-0.427***	-0.427***	-0.480***	-0.482***	-0.480***
	(0.053)	(0.051)	(0.061)	(0.058)	(0.061)
Citations	0.002***	0.002***	0.002***	0.002***	0.002***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Application experience					
nb projects	0.022	0.020	0.013	0.011	0.013
	(0.021)	(0.022)	(0.020)	(0.021)	(0.020)
Ongoing grant	-0.056	-0.053	-0.054**	-0.052**	-0.054**
	(0.047)	(0.047)	(0.049)	(0.051)	(0.050)
Refusal	-0.008***	-0.010***	0.005**	0.005	0.005**
	(0.094)	(0.098)	(0.091)	(0.103)	(0.092)
Lab variables					
Lab size			0.000	0.000	0.000
			(0.000)	(0.000)	(0.000)
Rank					
A+			0.036***	0.037***	0.036***
			(0.022)	(0.022)	(0.022)
B			-0.040*	-0.039*	-0.040*
			(0.049)	(0.051)	(0.049)
C			0.018	0.021	0.019
			(0.235)	(0.239)	(0.235)
Missing			0.003	0.002	0.002
			(0.088)	(0.088)	(0.088)
Current application					
Team size			0.010***	0.010***	0.010***
			(0.018)	(0.018)	(0.018)
Project duration			-0.001	-0.001	-0.001
			(0.002)	(0.002)	(0.002)
Private part.			-0.027**	-0.026**	-0.027**
			(0.041)	(0.042)	(0.041)
Cluster			0.441***	0.441***	0.440***
			(0.074)	(0.072)	(0.075)
Scientific field	YES	YES	YES	YES	YES
Program	YES	YES	YES	YES	YES
Observations	21,063	21,063	21,063	21,063	21,063
rho	-0.39	-0.32	-0.33	-0.27	-0.33
Prob chi2	0.008	0.03	0.01	0.07	0.02

Notes: Robust standard-errors in parentheses, clustered at the project level. The control variables Scientific field and Program are included in the model but the mean marginal effects are not reported. Prob chi2 represents the p-value associated with the Wald test of independence; Ho: The selection equation and the structural equation are independent.

Table 1.10 – Factors that influence the probability to submit a project for the PI (Mean marginal effects reported)

	(1)	(2)	(3)	(4)	(5)
Novelty		0.001**		0.001**	0.002***
		(0.005)		(0.005)	(0.006)
Interdisciplinarity	0.039***		0.038***		0.040***
	(0.030)		(0.030)		(0.030)
Individual variables					
Age	-0.002***	-0.002***	-0.002***	-0.002***	-0.002***
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Male	0.005***	0.005***	0.005***	0.005***	0.005***
	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)
Status					
Professor	0.048***	0.049***	0.048***	0.049***	0.048***
	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)
Associate Researcher	0.046***	0.047***	0.044***	0.045***	0.044***
	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)
Researcher	0.082***	0.082***	0.080***	0.080***	0.080***
	(0.023)	(0.023)	(0.023)	(0.023)	(0.023)
Citations	0.001***	0.001***	0.001***	0.001***	0.001***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Application experience					
nb projects	0.006***	0.007***	0.006***	0.007***	0.006***
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
Ongoing grant	-0.033***	-0.033***	-0.033***	-0.033***	-0.033***
	(0.029)	(0.029)	(0.029)	(0.029)	(0.029)
Refusal	0.065***	0.066***	0.065***	0.066***	0.064***
	(0.021)	(0.021)	(0.021)	(0.021)	(0.021)
Lab variables					
Lab size			-0.000	-0.000	-0.000
			(0.000)	(0.000)	(0.000)
Nb prev project lab	0.000***	0.000***	0.000***	0.000***	0.000***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Rank					
A+			0.003**	0.003*	0.003**
			(0.013)	(0.013)	(0.013)
B			-0.016***	-0.017***	-0.016***
			(0.024)	(0.024)	(0.024)
C			-0.016	-0.017	-0.016
			(0.103)	(0.102)	(0.102)
Missing			0.002	0.002	0.002
			(0.056)	(0.055)	(0.056)
Scientific field	YES	YES	YES	YES	YES
Year	YES	YES	YES	YES	YES
Observations	139,494	139,494	139,494	139,494	139,494

Notes: Robust standard-errors in parentheses, clustered at the individual or project level. The control variables Scientific field and Year are included in the model but the mean marginal effects are not reported.

Table 1.11 – Factors that influence the probability to receive a grant for the PI (Mean marginal effects reported)

	(1)	(2)	(3)	(4)	(5)
Novelty		-0.009**		-0.008**	-0.009***
		(0.012)		(0.014)	(0.014)
Interdisciplinarity	-0.059***		-0.078***		-0.081***
	(0.061)		(0.088)		(0.086)
Individual variables					
Age	-0.002***	-0.002***	-0.004***	-0.004***	-0.004***
	(0.016)	(0.017)	(0.022)	(0.024)	(0.022)
Male	0.016	0.015	0.018	0.017	0.019
	(0.029)	(0.030)	(0.037)	(0.038)	(0.036)
Status					
Professor	0.075	0.075	0.071	0.072	0.072
	(0.085)	(0.093)	(0.107)	(0.126)	(0.102)
Associate Researcher	0.077	0.076	0.070	0.070	0.070
	(0.084)	(0.091)	(0.101)	(0.118)	(0.096)
Researcher	0.137	0.139	0.124	0.126	0.125
	(0.138)	(0.150)	(0.164)	(0.196)	(0.156)
Application experience					
nb projects	0.017	0.017	0.012	0.012	0.013
	(0.018)	(0.019)	(0.020)	(0.022)	(0.020)
Ongoing grant	-0.045	-0.044	-0.049	-0.048	-0.049
	(0.071)	(0.075)	(0.093)	(0.103)	(0.091)
Refusal	0.002***	0.001***	0.011	0.010	0.011*
	(0.070)	(0.078)	(0.108)	(0.133)	(0.102)
Lab variables					
Lab size			0.000*	0.000*	0.000*
			(0.000)	(0.000)	(0.000)
Rank					
A+			0.043***	0.044***	0.043***
			(0.035)	(0.036)	(0.035)
B			-0.062*	-0.061*	-0.062*
			(0.081)	(0.088)	(0.080)
C			-0.049	-0.044	-0.047
			(0.275)	(0.282)	(0.273)
Missing			-0.018	-0.022	-0.021
			(0.148)	(0.151)	(0.146)
Current application					
Team size			0.012***	0.012***	0.012***
			(0.028)	(0.029)	(0.028)
Project duration			-0.001	-0.001	-0.001
			(0.003)	(0.003)	(0.003)
Private part.			-0.025	-0.024	-0.025
			(0.056)	(0.057)	(0.055)
Cluster			0.465***	0.465***	0.465***
			(0.143)	(0.148)	(0.142)
Citations	0.003*	0.003*	0.003***	0.003***	0.003***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Scientific field	YES	YES	YES	YES	YES
Year	YES	YES	YES	YES	YES
Program	YES	YES	YES	YES	YES
Observations	8,685	8,685	8,685	8,685	8,685
rho	-0.76	-0.72	-0.42	-0.36	-0.44
Prob chi2	0.000	0.000	0.049	0.17	0.03

Notes: Robust standard-errors in parentheses, clustered at the project level. The control variables Scientific field, Year and Program are included in the model but the mean marginal effects are not reported. Prob chi2 represents the p-value associated with the Wald test of independence; Ho: The selection equation and the structural equation are independent.

Table 1.12 – Factors that influence the probability to submit a project for the directed programs (Mean marginal effects reported)

	(1)	(2)	(3)	(4)	(5)
Novelty		0.004***		0.004***	0.005***
		(0.006)		(0.006)	(0.006)
Interdisciplinarity	0.068***		0.067***		0.071***
	(0.033)		(0.033)		(0.033)
Individual variables					
Age	-0.000***	-0.000***	-0.000***	-0.000***	-0.000***
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
Male	0.007***	0.009***	0.008***	0.009***	0.008***
	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)
Status					
Professor	0.056***	0.058***	0.056***	0.058***	0.056***
	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)
Associate Researcher	0.039***	0.040***	0.038***	0.039***	0.038***
	(0.019)	(0.019)	(0.019)	(0.019)	(0.019)
Researcher	0.083***	0.084***	0.082***	0.083***	0.082***
	(0.024)	(0.024)	(0.024)	(0.024)	(0.024)
Citations	0.001***	0.001***	0.001***	0.001***	0.001***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Application experience					
nb projects	0.023***	0.023***	0.023***	0.023***	0.022***
	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)
Ongoing grant	-0.012***	-0.013***	-0.013***	-0.013***	-0.013***
	(0.028)	(0.028)	(0.028)	(0.028)	(0.028)
Refusal	0.076***	0.078***	0.076***	0.078***	0.076***
	(0.023)	(0.023)	(0.023)	(0.023)	(0.023)
Lab variables					
Lab size			0.000	0.000	0.000
			(0.000)	(0.000)	(0.000)
Nb prev project lab	0.000**	0.000***	-0.000	0.000	-0.000
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Rank					
A+			-0.005**	-0.005***	-0.004**
			(0.015)	(0.015)	(0.015)
B			-0.020***	-0.022***	-0.020***
			(0.024)	(0.023)	(0.024)
C			-0.025*	-0.027**	-0.025*
			(0.107)	(0.105)	(0.107)
Missing			0.011*	0.012*	0.012*
			(0.056)	(0.056)	(0.056)
Scientific field	YES	YES	YES	YES	YES
Year	YES	YES	YES	YES	YES
Observations	138,066	138,066	138,066	138,066	138,066

Notes: Robust standard-errors in parentheses, clustered at the individual or project level. The control variables Scientific field and Year are included in the model but the mean marginal effects are not reported. The size of the sample is 138,066 since when a researcher apply to a directed program a given year, we deleted the other application to non-directed programs this same year if necessary.

Table 1.13 – Factors that influence the probability to receive a grant for the directed programs (Mean marginal effects reported)

	(1)	(2)	(3)	(4)	(5)
Novelty		-0.010***		-0.006**	-0.008***
		(0.011)		(0.013)	(0.013)
Interdisciplinarity	-0.113***		-0.085***		-0.089***
	(0.065)		(0.088)		(0.090)
Individual variables					
Age	-0.004***	-0.004***	-0.004***	-0.004***	-0.004***
	(0.019)	(0.023)	(0.022)	(0.024)	(0.022)
Male	0.012	0.011	0.004	0.003	0.004
	(0.027)	(0.030)	(0.032)	(0.033)	(0.032)
Status					
Professor	0.038***	0.036*	0.035	0.033	0.035
	(0.081)	(0.097)	(0.067)	(0.074)	(0.067)
Associate Researcher	0.063	0.061	0.063*	0.062*	0.063*
	(0.075)	(0.086)	(0.061)	(0.064)	(0.061)
Researcher	0.099*	0.097	0.091	0.092	0.092
	(0.127)	(0.149)	(0.094)	(0.103)	(0.094)
Citations	0.002	0.002	0.002***	0.002***	0.002***
	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)
Application experience					
nb projects	0.023**	0.021*	0.014	0.013	0.014
	(0.028)	(0.033)	(0.022)	(0.024)	(0.022)
Ongoing grant	-0.073	-0.069*	-0.067***	-0.065***	-0.067***
	(0.052)	(0.055)	(0.056)	(0.057)	(0.056)
Refusal	-0.029***	-0.031***	-0.008**	-0.009*	-0.008**
	(0.079)	(0.098)	(0.077)	(0.085)	(0.077)
Lab variables					
Lab size			0.000	0.000	0.000
			(0.000)	(0.000)	(0.000)
Rank					
A+			0.044***	0.044***	0.043***
			(0.029)	(0.029)	(0.029)
B			-0.025	-0.023	-0.026
			(0.059)	(0.061)	(0.059)
C			0.035	0.038	0.036
			(0.266)	(0.268)	(0.266)
Missing			0.040	0.040	0.039
			(0.108)	(0.109)	(0.108)
Current application					
Team size			0.003	0.003	0.003
			(0.023)	(0.024)	(0.023)
Project duration			-0.002**	-0.002**	-0.002**
			(0.003)	(0.003)	(0.003)
Private part.			-0.020	-0.019	-0.020
			(0.047)	(0.048)	(0.047)
Cluster			0.408***	0.407***	0.408***
			(0.074)	(0.073)	(0.074)
Scientific field	YES	YES	YES	YES	YES
Year	YES	YES	YES	YES	YES
Program	YES	YES	YES	YES	YES
Observations	10,687	10,687	10,687	10,687	10,687
rho	-0.72	-0.64	-0.32	-0.27	-0.32
Prob chi2	0.003	0.02	0.02	0.07	0.02

Notes: Robust standard-errors in parentheses, clustered at the project level. The control variables Scientific field, Year and Program are included in the model but the mean marginal effects are not reported. Prob chi2 represents the p-value associated with the Wald test of independence; Ho: The selection equation and the structural equation are independent.

Table 1.14 – Factors that influence the probability to submit a project for the non-directed program (Mean marginal effects reported)

	(1)	(2)	(3)	(4)	(5)
Novelty		-0.000		-0.000	0.001
		(0.005)		(0.005)	(0.005)
Interdisciplinarity	0.036***		0.035***		0.036***
	(0.033)		(0.033)		(0.034)
Individual variables					
Age	-0.002***	-0.002***	-0.002***	-0.002***	-0.002***
	(0.006)	(0.006)	(0.007)	(0.006)	(0.007)
Male	0.007***	0.008***	0.007***	0.008***	0.007***
	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)
Status					
Professor	0.047***	0.048***	0.046***	0.047***	0.046***
	(0.019)	(0.019)	(0.019)	(0.019)	(0.019)
Associate Researcher	0.049***	0.050***	0.047***	0.047***	0.047***
	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)
Researcher	0.082***	0.082***	0.079***	0.080***	0.079***
	(0.025)	(0.025)	(0.025)	(0.025)	(0.025)
Citations	0.001***	0.001***	0.001***	0.001***	0.001***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Application experience					
nb projects	0.015***	0.016***	0.015***	0.016***	0.015***
	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)
Ongoing grant	-0.038***	-0.038***	-0.038***	-0.038***	-0.038***
	(0.029)	(0.029)	(0.029)	(0.029)	(0.029)
Refusal	0.085***	0.086***	0.085***	0.086***	0.085***
	(0.022)	(0.022)	(0.022)	(0.022)	(0.022)
Lab variables					
Lab size			-0.000	-0.000*	-0.000
			(0.000)	(0.000)	(0.000)
Nb prev project lab	0.000***	0.000***	0.000***	0.000***	0.000***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Rank					
A+			0.007***	0.007***	0.007***
			(0.014)	(0.014)	(0.015)
B			-0.016***	-0.017***	-0.016***
			(0.025)	(0.025)	(0.025)
C			-0.060***	-0.061***	-0.060***
			(0.158)	(0.157)	(0.158)
Missing			-0.004	-0.003	-0.004
			(0.065)	(0.065)	(0.065)
Scientific field	YES	YES	YES	YES	YES
Year	YES	YES	YES	YES	YES
Observations	138,196	138,196	138,196	138,196	138,196

Notes: Robust standard-errors in parentheses, clustered at the individual or project level. The control variables Scientific field and Year are included in the model but the mean marginal effects are not reported. The size of the sample is 138,196 since when a researcher apply to a non-directed program a given year, we deleted the other application to directed programs this same year if necessary.

Table 1.15 – Factors that influence the probability to receive a grant for the non-directed program (Mean marginal effects reported)

	(1)	(2)	(3)	(4)	(5)
Novelty		-0.004		-0.004	-0.005*
		(0.011)		(0.013)	(0.012)
Interdisciplinarity	-0.027***		-0.035***		-0.036***
	(0.070)		(0.093)		(0.089)
Individual variables					
Age	-0.002***	-0.002***	-0.003***	-0.003***	-0.003***
	(0.017)	(0.019)	(0.021)	(0.027)	(0.020)
Male	0.014	0.013	0.015	0.014	0.015
	(0.036)	(0.038)	(0.045)	(0.053)	(0.044)
Status					
Professor	0.055	0.055	0.055	0.056	0.055
	(0.116)	(0.126)	(0.161)	(0.212)	(0.154)
Associate Researcher	0.052	0.052	0.050	0.051	0.050
	(0.115)	(0.126)	(0.157)	(0.207)	(0.150)
Researcher	0.089	0.090	0.087	0.089	0.088
	(0.185)	(0.202)	(0.256)	(0.339)	(0.244)
Citations	0.002	0.002	0.002	0.002	0.002
	(0.002)	(0.002)	(0.003)	(0.004)	(0.003)
Application experience					
nb projects	0.008**	0.007*	0.005	0.004	0.005
	(0.028)	(0.031)	(0.038)	(0.051)	(0.037)
Ongoing grant	-0.037	-0.037	-0.037	-0.037	-0.037
	(0.088)	(0.094)	(0.121)	(0.154)	(0.117)
Refusal	0.022***	0.020**	0.011	0.010	0.011*
	(0.131)	(0.149)	(0.201)	(0.285)	(0.189)
Lab variables					
Lab size			-0.000	0.000	-0.000
			(0.000)	(0.000)	(0.000)
Rank					
A+			0.019	0.020	0.019
			(0.043)	(0.050)	(0.042)
B			-0.048	-0.048	-0.048
			(0.104)	(0.125)	(0.101)
C			-0.074	-0.072	-0.072
			(0.525)	(0.591)	(0.513)
Missing			-0.043	-0.046	-0.044
			(0.159)	(0.170)	(0.156)
Current application					
Team size			0.009***	0.009***	0.009***
			(0.077)	(0.081)	(0.076)
Project duration			0.001	0.001	0.001
			(0.003)	(0.003)	(0.003)
Private part.			-0.017	-0.018	-0.017
			(0.075)	(0.080)	(0.074)
Cluster			0.738***	0.733***	0.740***
			(0.535)	(0.611)	(0.522)
Scientific field	YES	YES	YES	YES	YES
Year	YES	YES	YES	YES	YES
Observations	10,376	10,376	10,376	10,376	10,376
rho	-0.75	-0.69	-0.58	-0.46	-0.60
Prob chi2	0.01	0.03	0.14	0.37	0.11

Notes: Robust standard-errors in parentheses, clustered at the project level. The control variables Scientific field and Year are included in the model but the mean marginal effects are not reported. $Prob\,chi2$ *represents the p-value associated with the Wald test of independence; Ho: The selection equation and the structural equation are independent.*

Table 1.16 – Factors that influence the probability to receive a grant for the sample of teams (Mean marginal effects reported)

	(1)	(2)	(3)	(4)	(5)
Novelty team		-0.036***	-0.035***		-0.030***
		(0.036)	(0.036)		(0.037)
Interdisc. team	-0.099**		-0.091**	-0.081*	
	(0.138)		(0.138)	(0.142)	
Members diversity					
Field diversity	-0.072***	-0.070***	-0.064***	-0.075***	-0.074***
	(0.080)	(0.080)	(0.080)	(0.083)	(0.083)
Lab diversity	-0.075***	-0.078***	-0.077***	-0.075**	-0.078***
	(0.095)	(0.095)	(0.095)	(0.099)	(0.099)
Lab quality diversity	-0.055**	-0.055**	-0.053**	-0.064**	-0.064**
	(0.090)	(0.090)	(0.091)	(0.094)	(0.094)
Status diversity	-0.016	-0.014	-0.012	-0.025	-0.024
	(0.088)	(0.088)	(0.088)	(0.091)	(0.091)
Team variables					
Age team	-0.004***	-0.004**	-0.004**	-0.004**	-0.004**
	(0.043)	(0.043)	(0.043)	(0.044)	(0.044)
PI male	0.010	0.011	0.010	-0.002	-0.001
	(0.047)	(0.047)	(0.047)	(0.049)	(0.049)
Citations team	0.029***	0.030***	0.032***	0.030***	0.030***
	(0.022)	(0.022)	(0.023)	(0.023)	(0.024)
Top lab	0.061***	0.062***	0.060***	0.061***	0.062***
	(0.045)	(0.045)	(0.045)	(0.047)	(0.046)
Nb grants team	0.016***	0.017***	0.017***	0.017***	0.017***
	(0.019)	(0.020)	(0.020)	(0.020)	(0.020)
Team size	0.003	0.004	0.004	0.001	0.002
	(0.033)	(0.033)	(0.033)	(0.034)	(0.034)
Project duration	-0.001	-0.001	-0.001	-0.001	-0.001
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Private part.	-0.012	-0.015	-0.015	-0.010	-0.012
	(0.062)	(0.062)	(0.062)	(0.075)	(0.074)
Cluster	0.455***	0.454***	0.453***	0.448***	0.448***
	(0.083)	(0.083)	(0.083)	(0.086)	(0.086)
Experimental develop.				-0.040	-0.041
				(0.149)	(0.149)
Industrial research				-0.027	-0.029
				(0.088)	(0.088)
Programs	YES	YES	YES	YES	YES
Year	YES	YES	YES	YES	YES
Observations	5,057	5,057	5,057	4,670	4,670

Notes: Robust standard-errors in parentheses. The control variables Programs and Year are included in the model but the mean marginal effects are not reported. The type of research of the project is only known for 4,670 projects.

Appendix A. The disambiguation process

In this appendix, we present briefly the process followed to disambiguate authors' names reported on the scientific publications extracted from the WoS[25]. Based on our list of researchers and professors' surnames and first name initials, we were first able to extract a set of 7,7 billion publications from the WoS. We eventually admit that the author of the retrieved documents is indeed one of the researchers in our list if several conditions are verified. We developed two steps of validation, called "seed" and "expand", which rely on different subsets of conditions.

The SEED step

This step is composed of four conditions that need to be validated together:

- The surname and the initials of the first name of a researcher are identified among one of the authors' identities of a WoS document.
- The date of birth of a researcher is consistent with the year of publication of the document. We assume the publication activity of the researcher may start from 24 and we defined the upper bound at 80.
- The scientific field of a research (which is inferred from his known specialty) is consistent with one of the scientific disciplines associated with the journal[26] which publishes the document.
- The University or Public Research Organization which is the employer of a researcher is mentioned among the authors' addresses reported on the document.

The EXPAND step

The rejected documents in the "seed" step, that is to say the ones that do not validate jointly the previous four conditions, are examined again in the "expand" step. We remove the fourth condition of the previous step in order to reach documents published when a researcher was employed by another University, and we substitute it with one of the three extra conditions presented below. A rejected document is finally retained if we are able to observe some similarity with at least one of the documents selected in the "seed" step for the same author's surname.

In this second step we consider three specific information reported on the documents:

[25] A more detailed presentation is introduced in chapter 2, Appendix G, p.160. The disambiguation approach implemented with this database differs only in terms of the type of keywords used, from the one presented in chapter 2.

[26] The scientific disciplines of a journal are inferred from the WoS subject categories. We use the disciplines categorization implemented by the *Observatoire des Sciences et Techniques* (OST).

- The keywords (reported by the authors)
- The identities of the coauthors
- The list of references

The underlying idea considers that a research is likely to choose the same keywords and to cite the same references when he publishes different but related documents. Furthermore, he is likely to repeat the collaboration with some coauthors during is career.

A rejected document is finally retained in this second step if the first three conditions of the "seed" step are validated, as well as one of the following proposals:

- The rejected document is associated with a (not too highly used) keyword or/and with a coauthor which is also reported on a document validated in the "seed" step.
- The rejected document includes a high share of references that are also mentioned on a document validated in the "seed" step. For this condition, we estimate a similarity score between two documents which is based on shared references and controls for the number of cites each reference obtained.

Benchmarking

At the end of the disambiguation process, we are left with 1,1 billion validated articles that is to say approximately 14% of the initial stock of articles retrieved from the WoS. We assess the reliability of our disambiguation method based on a sample of 353 researchers and professors who verify themselves their list of publications (See chapter 2, Appendix G, p.164 for detailed information). We get a recall of 0.85 and a precision of 0.83[27]. So, we are able to recover about 85% of the true publications of a researcher on average, and among the full set of articles that we attribute to him/her, about 83% are well and truly his/hers .

[27]These two indicators are scored by: $RECALL = \frac{number\,of\,true\,positives}{number\,of\,true\,positives + number\,of\,false\,negatives}$ and $PRECISION = \frac{number\,of\,true\,positives}{number\,of\,true\,positives + number\,of\,false\,positives}$

Appendix B. Construction of the variables

In this appendix, we describe how the main variables of the study are constructed.

Scientific production

- Factional count of (3-year) citations of articles published by the researcher i in the three-year window preceding the application date t

$$Citations_i^{t-3\_t} = \frac{1}{Ni^{t-3\_t}} \sum_{t-3}^{t} \sum_{j \in J_{i,t}} \frac{CIT\,(j)}{n\,(j)}. \tag{1.3}$$

  with $CIT\,(j)$ the number of citations received by j and Ni^{t-3_t} the number of articles published by i between $t-3$ and t. $n\,(j)$ represents the number of authors of the article j.

- Highest Novelty index associated with an article published in the three-year period preceding the application date.

$$Novelty_i^{t-3\_t} = \max_{t-3}^{t} -log(10\,percentile(Com_{klct})) \tag{1.4}$$

  with $Com_{klct} = \frac{N_{klct} \times N_{ct}}{N_{kct} \times N_{lct}}$ the frequency of the combination of the keywords k and l, in the specialty c (Wos subject category) of the revue which publishes the article in t. N_{ct} is the number of keywords combinations of articles published in the specialty c in t and N_{klct} is the number of articles which report the keywords combination k,l in the specialty c in t (see Carayol et al. 2018 for more details).

- Mean degree of interdisciplinarity per article published in the three-year period preceding the application date (Simpson variety index)

$$Interdisc_i^{t-3\_t} = \frac{1}{Ni^{t-3\_t}} \sum_{t-3}^{t} 1 - Simp_j \tag{1.5}$$

  with $Simp_j = \sum_{dj} p_{dj}^2$ the Simpson index, where p_{dj} is the share of the discipline d in the total of the disciplines (deduced from the WoS subject categories associated with the journals) of the articles cited by j. We also calculate this indicator relying on subject categories instead of the research field.

Heterogeneity of the team

- Diversity of the scientific fields in the team.

$$Field\_diversity_e = 1 - \sum_{disc,e} p^2_{disc,e} \tag{1.6}$$

with *disc* the discipline of the members of the team e.

- Diversity of the type of laboratories of the members.

$$Lab\_diversity_e = 1 - \sum_{lab,e} p^2_{lab,e} \tag{1.7}$$

with *lab* the type of the laboratories of the members of the team e. The type of laboratories are : Under the responsibility of the "CNRS", or another public research organization apart from the CNRS, or the ministry of Higher Education and Research only.

- Diversity of the quality of the laboratories of the members.

$$Quality\_lab\_diversity_e = 1 - \sum_{score,e} p^2_{score,e} \tag{1.8}$$

with *score* the grade assigned to the laboratories of the members of the team e. The scores are in decreasing order: "A+", "A", "B" or "C".

- Diversity of the status of the members.

$$Status\_diversity_e = 1 - \sum_{status,e} p^2_{status,e} \tag{1.9}$$

with *status* the status of the members of the team e. The status are "Assistant Professor", "Full professor", "Assistant Researcher" or "Research director".

Additional variables for the robustness analysis

- Mean degree of interdisciplinarity per article published in the three-year window preceding the application date (Shannon index).

$$Interdisc\_Shannon_i^{t-3\_t} = \frac{1}{Ni^{t-3\_t}} \sum_{t-3}^{t} Shannon_j \tag{1.10}$$

with $Shannon_j = -\sum_{dj} p_{dj}\ log\ p_{dj}$ the Shannon index, where p_{dj} is the share of the discipline d in the total of the disciplines (deduced from the WoS subject categories associated with the journals) of the articles cited by j. We also calculate this indicator relying on subject categories instead of the research field.

- Fractional count of articles published in the three-year window preceding the application date

$$Volume_i^{t-3\_t} = \sum_{t-3}^{t} \sum_{j \in J_{i,t}} \frac{1}{n\left(j\right)}, \tag{1.11}$$

with $J_{i,t}$ the set of articles published by i in t.

- Highest Impact Factor of a journal which publishes an article in the three-year period preceding the application date

$$IF_i^{t-3\_t} = \max_{t-3}^{t} IF_{j,t} \tag{1.12}$$

with $IF_{j,t}$ the Impact Factor of the revue which publishes the article j in t.

Appendix C. Robustness Check estimations.

Table 1.17 – Robustness Check 1: Factors that influence the probability to submit a project for the whole sample (Mean marginal effects reported)

	(1)	(2)	(3)
Interdisc. Simpson (SC)	0.066***		
	(0.019)		
Interdisc. Shannon (field)		0.056***	
		(0.017)	
Interdisc. Shannon (SC)			0.030***
			(0.009)
Individual variables			
Age	-0.001***	-0.001***	-0.001***
	(0.006)	(0.006)	(0.006)
Male	0.012***	0.013***	0.012***
	(0.013)	(0.013)	(0.013)
Status			
Professor	0.092***	0.094***	0.092***
	(0.016)	(0.016)	(0.016)
Associate Researcher	0.075***	0.077***	0.075***
	(0.016)	(0.016)	(0.016)
Researcher	0.150***	0.152***	0.150***
	(0.022)	(0.022)	(0.022)
Citations	0.001***	0.002***	0.001***
	(0.001)	(0.001)	(0.001)
Application experience			
nb projects	0.042***	0.042***	0.042***
	(0.009)	(0.009)	(0.009)
Ongoing grant	-0.055***	-0.055***	-0.055***
	(0.024)	(0.024)	(0.024)
Refusal	0.164***	0.164***	0.164***
	(0.019)	(0.019)	(0.019)
Lab variables			
Lab size	-0.000	0.000	-0.000
	(0.000)	(0.000)	(0.000)
Nb prev project lab	0.000**	0.000**	0.000**
	(0.001)	(0.001)	(0.001)
Rank			
A+	0.002	0.003	0.002
	(0.013)	(0.013)	(0.013)
B	-0.033***	-0.033***	-0.034***
	(0.020)	(0.020)	(0.020)
C	-0.061***	-0.062***	-0.062***
	(0.100)	(0.100)	(0.100)
Missing	0.009	0.011	0.009
	(0.051)	(0.051)	(0.051)
Scientific field	YES	YES	YES
Year	YES	YES	YES
Observations	139,494	139,494	139,494

Notes: Robust standard-errors in parentheses, clustered at the individual or project level. The control variables Scientific field and Year are included in the model but the mean marginal effects are not reported. "SC" refers to the interdisciplinarity measures based on the WoS "subject categories", whereas "field" uses the scientific field.

Table 1.18 – Robustness Check 1: Factors that influence the probability to receive a grant for the whole sample (Mean marginal effects reported)

	(1)	(2)	(3)
Interdisc. Simpson (SC)	-0.020***		
	(0.052)		
Interdisc. Shannon (field)		-0.041***	
		(0.041)	
Interdisc. Shannon (SC)			-0.012***
			(0.022)
Individual variables			
Age	-0.003***	-0.003***	-0.003***
	(0.017)	(0.017)	(0.017)
Male	0.009	0.010	0.009
	(0.025)	(0.025)	(0.025)
Status			
Professor	0.048	0.048	0.048
	(0.066)	(0.065)	(0.065)
Associate Researcher	0.057	0.056	0.057
	(0.059)	(0.058)	(0.058)
Researcher	0.095	0.094	0.095
	(0.099)	(0.098)	(0.097)
Citations	0.002***	0.002***	0.002***
	(0.001)	(0.001)	(0.001)
Application experience			
nb projects	0.012	0.013	0.012
	(0.020)	(0.020)	(0.020)
Ongoing grant	-0.053**	-0.054**	-0.053**
	(0.050)	(0.049)	(0.050)
Refusal	0.004**	0.005**	0.004**
	(0.095)	(0.091)	(0.092)
Lab variables			
Lab size	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)
Rank			
A+	0.037***	0.036***	0.037***
	(0.022)	(0.022)	(0.022)
B	-0.038*	-0.039*	-0.038*
	(0.050)	(0.049)	(0.049)
C	0.019	0.019	0.020
	(0.236)	(0.235)	(0.235)
Missing	0.003	0.003	0.004
	(0.088)	(0.088)	(0.087)
Current application			
Team size	0.010***	0.010***	0.010***
	(0.018)	(0.018)	(0.018)
Project duration	-0.001	-0.001	-0.001
	(0.002)	(0.002)	(0.002)
Private part.	-0.027**	-0.027**	-0.027**
	(0.042)	(0.041)	(0.041)
Cluster	0.441***	0.441***	0.441***
	(0.074)	(0.075)	(0.075)
Scientific field	YES	YES	YES
Year	YES	YES	YES
Program	YES	YES	YES
Observations	21,063	21,063	21,063
rho	-0.33	-0.33	-0.27
Prob chi2	0.01	0.01	0.07

Notes: Robust standard-errors in parentheses, clustered at the project level. The control variables Scientific field and Year are included in the model but the mean marginal effects are not reported. "SC" refers to the interdisciplinarity measures based on the WoS "subject categories", whereas "field" uses the scientific field. Prob chi2 represents the p-value associated with the Wald test of independence; Ho: The selection equation and the structural equation are independent.

Table 1.19 – Robustness Check 2: Factors that influence the probability to submit a project for the whole sample (Mean marginal effects reported)

	(1)	(2)	(3)	(4)
Novelty		0.003***		0.003***
		(0.005)		(0.005)
Interdisciplinarity	0.068***		0.077***	
	(0.029)		(0.029)	
Individual variables				
Age	-0.001***	-0.001***	-0.001***	-0.001***
	(0.006)	(0.006)	(0.006)	(0.006)
Male	0.010***	0.011***	0.012***	0.012***
	(0.013)	(0.013)	(0.013)	(0.013)
Status				
Professor	0.087***	0.087***	0.088***	0.089***
	(0.016)	(0.016)	(0.016)	(0.016)
Associate Researcher	0.072***	0.072***	0.076***	0.076***
	(0.016)	(0.016)	(0.016)	(0.016)
Researcher	0.143***	0.141***	0.148***	0.147***
	(0.022)	(0.022)	(0.022)	(0.022)
Volume	0.011***	0.013***	0.013***	0.016***
	(0.005)	(0.005)	(0.005)	(0.005)
IF	0.002***	0.002***		
	(0.001)	(0.001)		
Nb hit 10%			0.003**	0.003**
			(0.007)	(0.007)
Application experience				
nb projects	0.041***	0.041***	0.042***	0.042***
	(0.009)	(0.009)	(0.009)	(0.009)
Ongoing grant	-0.056***	-0.057***	-0.055***	-0.055***
	(0.024)	(0.024)	(0.024)	(0.024)
Refusal	0.164***	0.165***	0.164***	0.166***
	(0.019)	(0.019)	(0.019)	(0.019)
Lab variables				
Lab size	0.000	-0.000	0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Nb prev project lab	0.000**	0.000***	0.000**	0.000***
	(0.001)	(0.001)	(0.001)	(0.001)
Rank				
A+	0.001	0.001	0.003	0.002
	(0.013)	(0.013)	(0.013)	(0.013)
B	-0.031***	-0.032***	-0.032***	-0.033***
	(0.020)	(0.020)	(0.020)	(0.020)
C	-0.059***	-0.060***	-0.061***	-0.063***
	(0.100)	(0.098)	(0.100)	(0.099)
Missing	0.009	0.010	0.010	0.010
	(0.051)	(0.051)	(0.051)	(0.051)
Scientific field	YES	YES	YES	YES
Year	YES	YES	YES	YES
Observations	139,494	139,494	139,494	139,494

Notes: Robust standard-errors in parentheses, clustered at the individual or project level. The control variables Scientific field and Year are included in the model but the mean marginal effects are not reported.

Table 1.20 – Robustness Check 2: Factors that influence the probability to receive a grant for the whole sample (Mean marginal effects reported)

	(1)	(2)	(3)	(4)	(5)	(6)
Novelty		-0.006**		-0.005**		-0.006***
		(0.009)		(0.009)		(0.009)
Interdisciplinarity	-0.092***		-0.082***		-0.066***	
	(0.064)		(0.066)		(0.072)	
Individual variables						
Age	-0.003***	-0.003***	-0.003***	-0.003***	-0.003***	-0.003***
	(0.017)	(0.018)	(0.017)	(0.018)	(0.018)	(0.020)
Male	0.007	0.007	0.010	0.010	0.007	0.006
	(0.025)	(0.025)	(0.025)	(0.025)	(0.026)	(0.027)
Status						
Professor	0.045	0.046	0.047	0.049	0.053	0.053
	(0.060)	(0.063)	(0.063)	(0.067)	(0.076)	(0.082)
Associate Researcher	0.052*	0.053**	0.056*	0.057**	0.060	0.060*
	(0.054)	(0.056)	(0.057)	(0.061)	(0.067)	(0.071)
Researcher	0.087*	0.091**	0.096	0.099**	0.098	0.100
	(0.087)	(0.092)	(0.095)	(0.102)	(0.112)	(0.120)
Volume	0.009**	0.007	0.012**	0.010*		
	(0.009)	(0.010)	(0.010)	(0.011)		
IF	0.003***	0.003***				
	(0.001)	(0.002)				
Nb hit 10%			0.007*	0.007		
			(0.012)	(0.012)		
Citations					0.002***	0.002***
					(0.001)	(0.001)
Application experience						
nb projects	0.011	0.010	0.013	0.012	0.012	0.011
	(0.018)	(0.019)	(0.019)	(0.021)	(0.023)	(0.025)
Ongoing grant	-0.053**	-0.051***	-0.051**	-0.048**	-0.049	-0.047*
	(0.049)	(0.050)	(0.049)	(0.050)	(0.055)	(0.056)
Refusal	0.006	0.006	0.004*	0.004	0.009*	0.009
	(0.088)	(0.097)	(0.091)	(0.102)	(0.107)	(0.119)
Lab variables						
Lab size	0.000	0.000	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Rank						
A+	0.035***	0.036***	0.038***	0.038***	0.035***	0.036***
	(0.022)	(0.022)	(0.022)	(0.022)	(0.023)	(0.023)
B	-0.040**	-0.039**	-0.040**	-0.039**	-0.045**	-0.044**
	(0.048)	(0.050)	(0.049)	(0.051)	(0.052)	(0.054)
C	0.026	0.030	0.014	0.018	-0.002	0.002
	(0.238)	(0.241)	(0.238)	(0.242)	(0.265)	(0.269)
Missing	0.002	0.001	0.000	-0.000	-0.004	-0.006
	(0.089)	(0.090)	(0.088)	(0.089)	(0.092)	(0.093)
Current application						
Team size	0.011***	0.011***	0.010***	0.010***	0.008***	0.008***
	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)
Project duration	-0.001	-0.001	-0.001	-0.001	-0.000	-0.000
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Private part.	-0.024*	-0.024*	-0.027**	-0.027**	-0.029*	-0.029*
	(0.042)	(0.043)	(0.042)	(0.042)	(0.050)	(0.051)
Cluster	0.441***	0.441***	0.442***	0.443***	0.435***	0.435***
	(0.068)	(0.065)	(0.072)	(0.069)	(0.081)	(0.077)
Experimental develop.					-0.043	-0.043
					(0.093)	(0.095)
Industrial research					-0.015	-0.015
					(0.059)	(0.060)
Scientific field	YES	YES	YES	YES	YES	YES
Year	YES	YES	YES	YES	YES	YES
Program	YES	YES	YES	YES	YES	YES
Observations	21,063	21,063	21,063	21,063	19,471	19,471
rho	-0.25	-0.18	-0.29	-0.23	-0.33	-0.26
Prob chi2	0.046	0.19	0.03	0.12	0.04	0.13

Notes: Robust standard-errors in parentheses, clustered at the project level. The control variables Scientific field, Year and Program are included in the model but the mean marginal effects are not reported. Prob chi2 represents the p-value associated with the Wald test of independence; Ho: The selection equation and the structural equation are independent. The type of research of the project is only known for 19,471 partnerxprojects.

Table 1.21 – Robustness Check 3: Factors that influence the probability to receive a grant for the whole sample (Mean marginal effects reported)

	(1)	(2)	(3)	(4)
Novelty		-0.005**		-0.005**
		(0.010)		(0.010)
Interdisciplinarity	-0.043***		-0.045***	
	(0.072)		(0.073)	
Individual variables				
Age	-0.003***	-0.003***	-0.003***	-0.003***
	(0.017)	(0.018)	(0.017)	(0.018)
Male	0.008	0.008	0.008	0.008
	(0.028)	(0.029)	(0.028)	(0.029)
Status				
Professor	0.040	0.040	0.041	0.042
	(0.063)	(0.068)	(0.064)	(0.070)
Associate Researcher	0.053	0.054*	0.054	0.054*
	(0.058)	(0.062)	(0.059)	(0.063)
Researcher	0.083	0.084	0.083	0.085
	(0.092)	(0.099)	(0.093)	(0.101)
Citations	0.002***	0.002***	0.002***	0.002***
	(0.001)	(0.001)	(0.001)	(0.001)
Application experience				
nb projects	0.009	0.008	0.009	0.008
	(0.020)	(0.022)	(0.020)	(0.022)
Ongoing grant	-0.043	-0.042*	-0.042	-0.041*
	(0.052)	(0.053)	(0.052)	(0.054)
Refusal	0.008***	0.008**	0.008***	0.008*
	(0.087)	(0.097)	(0.088)	(0.099)
Lab variables				
Lab size	-0.000	-0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Rank				
A+	0.022***	0.022***	0.022***	0.022***
	(0.024)	(0.024)	(0.024)	(0.024)
B	-0.036	-0.035*	-0.035	-0.035*
	(0.053)	(0.055)	(0.054)	(0.056)
C	0.056	0.059	0.048	0.051
	(0.299)	(0.304)	(0.317)	(0.322)
Missing	-0.034	-0.036	-0.039	-0.041*
	(0.101)	(0.102)	(0.104)	(0.105)
Current application				
Team size	0.008***	0.008***	0.008***	0.007***
	(0.023)	(0.023)	(0.023)	(0.023)
Project duration	0.000	0.000	0.000	0.000
	(0.003)	(0.003)	(0.003)	(0.003)
Private part.	0.016	0.016	0.016	0.016
	(0.051)	(0.051)	(0.051)	(0.052)
Cluster	0.398***	0.398***	0.404***	0.404***
	(0.090)	(0.088)	(0.092)	(0.090)
Program Call <0.8%	YES	YES	NO	NO
Program Call <0.6%	NO	NO	YES	YES
Scientific field	YES	YES	YES	YES
Year	YES	YES	YES	YES
Specific Program	YES	YES	YES	YES
Observations	18,940	18,940	18,843	18,843
rho	-0.40	-0.35	-0.40	-0.34
Prob chi2	0.001	0.014	0.002	0.02

Notes: Robust standard-errors in parentheses, clustered at the project level. The control variables Scientific field, Year and dummies for the specific program call are included in the model but the mean marginal effects are not reported. Prob chi2 represents the p-value associated with the Wald test of independence; Ho: The selection equation and the structural equation are independent.

Appendix D. Estimated coefficients associated with the Heckman probit selection models and the standard probit model.

Table 1.22 – Factors that influence the probability to submit a project for the whole sample (Heckman Probit coefficients reported)

	(1)	(2)	(3)	(4)	(5)
Novelty		0.016***		0.017***	0.030***
		(0.005)		(0.005)	(0.005)
Interdisciplinarity	0.506***		0.500***		0.524***
	(0.028)		(0.028)		(0.029)
Individual variables					
Age	0.127***	0.129***	0.128***	0.130***	0.128***
	(0.006)	(0.005)	(0.006)	(0.006)	(0.006)
Age squared	-0.002***	-0.002***	-0.002***	-0.002***	-0.002***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Male	0.071***	0.080***	0.072***	0.081***	0.072***
	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)
Status					
Professor	0.513***	0.526***	0.512***	0.524***	0.511***
	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)
Associate Researcher	0.434***	0.436***	0.419***	0.420***	0.419***
	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)
Researcher	0.847***	0.849***	0.830***	0.833***	0.828***
	(0.022)	(0.022)	(0.022)	(0.022)	(0.022)
Year					
2006	0.482***	0.487***	0.487***	0.491***	0.487***
	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)
2007	0.215***	0.224***	0.227***	0.233***	0.228***
	(0.021)	(0.021)	(0.021)	(0.021)	(0.021)
2008	0.019	0.032	0.037	0.046*	0.038
	(0.023)	(0.023)	(0.023)	(0.023)	(0.023)
2009	0.239***	0.254***	0.262***	0.272***	0.263***
	(0.022)	(0.022)	(0.023)	(0.023)	(0.023)
Citations	0.009***	0.009***	0.008***	0.009***	0.008***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Application experience					
nb projects	0.227***	0.233***	0.227***	0.233***	0.226***
	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)
Ongoing grant	-0.296***	-0.299***	-0.300***	-0.302***	-0.300***
	(0.024)	(0.024)	(0.024)	(0.024)	(0.024)
Refusal	0.894***	0.904***	0.894***	0.904***	0.893***
	(0.019)	(0.019)	(0.019)	(0.019)	(0.019)
Lab variables					
Lab size			0.000	-0.000	0.000
			(0.000)	(0.000)	(0.000)
Nb prev project lab	0.003***	0.003***	0.002**	0.002***	0.001**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Rank					
A+			0.014	0.011	0.017
			(0.013)	(0.013)	(0.013)
B			-0.177***	-0.189***	-0.176***
			(0.020)	(0.020)	(0.020)
C			-0.336***	-0.351***	-0.337***
			(0.100)	(0.099)	(0.100)
Missing			0.060	0.064	0.064
			(0.051)	(0.051)	(0.051)
Scientific field	YES	YES	YES	YES	YES
Observations	139,494	139,494	139,494	139,494	139,494

Notes: Robust standard-errors in parentheses, clustered at the individual or project level. The control variables Scientific field is included in the model but the coefficients are not reported.

Table 1.23 – Factors that influence the probability to receive a grant for the whole sample (Heckman Probit coefficients reported)

	(1)	(2)	(3)	(4)	(5)
Novelty		-0.029***		-0.025***	-0.031***
		(0.009)		(0.009)	(0.009)
Interdisciplinarity	-0.356***		-0.342***		-0.357***
	(0.067)		(0.068)		(0.070)
Individual variables					
Age	-0.060***	-0.054***	-0.070***	-0.065***	-0.070***
	(0.017)	(0.018)	(0.017)	(0.018)	(0.017)
Age squared	0.001***	0.001***	0.001***	0.001***	0.001***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Male	0.018	0.019	0.015	0.015	0.016
	(0.024)	(0.024)	(0.025)	(0.026)	(0.025)
Status					
Professor	-0.012	0.017	0.041	0.063	0.041
	(0.069)	(0.071)	(0.065)	(0.072)	(0.066)
Associate Researcher	0.028	0.056	0.092	0.113*	0.092
	(0.063)	(0.063)	(0.058)	(0.063)	(0.059)
Researcher	0.023	0.080	0.119	0.164	0.120
	(0.106)	(0.107)	(0.097)	(0.106)	(0.099)
Year					
2006	-1.004***	-1.000***	-1.133***	-1.127***	-1.132***
	(0.049)	(0.053)	(0.053)	(0.057)	(0.053)
2007	-0.942***	-0.953***	-1.332***	-1.342***	-1.330***
	(0.051)	(0.050)	(0.060)	(0.058)	(0.060)
2008	-0.839***	-0.864***	-1.187***	-1.206***	-1.185***
	(0.059)	(0.055)	(0.066)	(0.064)	(0.066)
2009	-1.296***	-1.316***	-1.652***	-1.669***	-1.651***
	(0.053)	(0.051)	(0.060)	(0.058)	(0.060)
Citations	0.004***	0.004***	0.005***	0.005***	0.005***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Application experience					
nb projects	-0.011	-0.003	-0.012	-0.006	-0.011
	(0.021)	(0.022)	(0.020)	(0.021)	(0.020)
Ongoing grant	-0.064	-0.076	-0.107**	-0.116**	-0.108**
	(0.047)	(0.047)	(0.049)	(0.051)	(0.050)
Refusal	-0.309***	-0.265***	-0.188**	-0.152	-0.188**
	(0.094)	(0.098)	(0.091)	(0.102)	(0.093)
Lab variables					
Lab size			0.000	0.000	0.000
			(0.000)	(0.000)	(0.000)
Rank					
A+			0.116***	0.119***	0.113***
			(0.022)	(0.022)	(0.022)
B			-0.093*	-0.097*	-0.095*
			(0.049)	(0.051)	(0.049)
C			0.144	0.144	0.148
			(0.236)	(0.239)	(0.235)
Missing			-0.006	-0.009	-0.011
			(0.088)	(0.088)	(0.088)
Current application					
Team size			0.067***	0.066***	0.067***
			(0.017)	(0.017)	(0.017)
Team size squared			-0.005***	-0.005***	-0.005***
			(0.001)	(0.001)	(0.001)
Private part.			0.706***	0.746***	0.712***
			(0.242)	(0.247)	(0.242)
Project duration			-0.002	-0.002	-0.002
			(0.002)	(0.002)	(0.002)
Cluster			1.443***	1.463***	1.441***
			(0.073)	(0.071)	(0.074)
Scientific field	YES	YES	YES	YES	YES
Program	YES	YES	YES	YES	YES
Observations	21,063	21,063	21,063	21,063	21,063
rho	-0.39	-0.32	-0.32	-0.26	-0.32
Prob chi2	0.008	0.03	0.02	0.08	0.02

Notes: Robust standard-errors in parentheses, clustered at the project level. The control variables Scientific field and Program are included in the model but the coefficients are not reported. Prob chi2 represents the p-value associated with the Wald test of independence; Ho: The selection equation and the structural equation are independent.

Table 1.24 – Factors that influence the probability to submit a project for the PI (Heckman Probit coefficients reported)

	(1)	(2)	(3)	(4)	(5)
Novelty		0.011**		0.012**	0.020***
		(0.005)		(0.005)	(0.006)
Interdisciplinarity	0.360***		0.355***		0.370***
	(0.030)		(0.030)		(0.030)
Individual variables					
Age	0.080***	0.082***	0.081***	0.083***	0.081***
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Age squared	-0.001***	-0.001***	-0.001***	-0.001***	-0.001***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Male	0.043***	0.050***	0.044***	0.051***	0.044***
	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)
Status					
Professor	0.446***	0.454***	0.444***	0.452***	0.443***
	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)
Associate Researcher	0.430***	0.432***	0.414***	0.416***	0.414***
	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)
Researcher	0.761***	0.762***	0.745***	0.747***	0.743***
	(0.023)	(0.023)	(0.023)	(0.023)	(0.023)
Citations	0.006***	0.007***	0.006***	0.007***	0.006***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Application experience					
nb projects	0.059***	0.063***	0.058***	0.062***	0.057***
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
Ongoing grant	-0.303***	-0.305***	-0.306***	-0.308***	-0.306***
	(0.029)	(0.029)	(0.029)	(0.029)	(0.029)
Refusal	0.602***	0.611***	0.602***	0.611***	0.601***
	(0.021)	(0.021)	(0.021)	(0.021)	(0.021)
Lab variables					
Lab size			-0.000	-0.000	-0.000
			(0.000)	(0.000)	(0.000)
Nb prev project lab	0.003***	0.003***	0.003***	0.003***	0.002***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Rank					
A+			0.028**	0.025*	0.030**
			(0.013)	(0.013)	(0.013)
B			-0.150***	-0.157***	-0.149***
			(0.024)	(0.024)	(0.024)
C			-0.146	-0.159	-0.147
			(0.103)	(0.102)	(0.102)
Missing			0.016	0.020	0.019
			(0.056)	(0.055)	(0.056)
Scientific field	YES	YES	YES	YES	YES
Year	YES	YES	YES	YES	YES
Observations	139,494	139,494	139,494	139,494	139,494

Notes: Robust standard-errors in parentheses, clustered at the individual or project level. The control variables Scientific field and Year are included in the model but the coefficients are not reported.

Table 1.25 – Factors that influence the probability to receive a grant for the PI (Heckman Probit coefficients reported)

	(1)	(2)	(3)	(4)	(5)
Novelty		-0.029**		-0.030**	-0.036***
		(0.012)		(0.014)	(0.014)
Interdisciplinarity	-0.380***		-0.376***		-0.394***
	(0.061)		(0.087)		(0.086)
Individual variables					
Age	-0.060***	-0.060***	-0.077***	-0.075***	-0.079***
	(0.016)	(0.017)	(0.022)	(0.024)	(0.022)
Age squared	0.001***	0.001***	0.001***	0.001***	0.001***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Male	0.008	0.006	0.044	0.043	0.044
	(0.029)	(0.030)	(0.037)	(0.038)	(0.036)
Status					
Professor	-0.123	-0.102	0.075	0.103	0.068
	(0.085)	(0.093)	(0.106)	(0.125)	(0.102)
Associate Researcher	-0.107	-0.083	0.081	0.109	0.073
	(0.084)	(0.091)	(0.101)	(0.117)	(0.097)
Researcher	-0.185	-0.138	0.140	0.197	0.127
	(0.138)	(0.150)	(0.164)	(0.195)	(0.157)
Citations	0.003*	0.003*	0.007***	0.007***	0.007***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Application experience					
nb projects	0.000	0.001	0.019	0.021	0.019
	(0.018)	(0.019)	(0.020)	(0.022)	(0.020)
Ongoing grant	0.096	0.083	-0.049	-0.066	-0.043
	(0.071)	(0.075)	(0.093)	(0.102)	(0.091)
Refusal	-0.399***	-0.386***	-0.177*	-0.151	-0.188*
	(0.070)	(0.078)	(0.107)	(0.131)	(0.102)
Lab variables					
Lab size			0.001*	0.001*	0.001*
			(0.000)	(0.000)	(0.000)
Rank					
A+			0.132***	0.138***	0.128***
			(0.035)	(0.036)	(0.035)
B			-0.154*	-0.163*	-0.149*
			(0.081)	(0.088)	(0.080)
C			-0.104	-0.096	-0.095
			(0.275)	(0.282)	(0.273)
Missing			-0.069	-0.083	-0.079
			(0.148)	(0.151)	(0.146)
Current application					
Team size			0.110***	0.110***	0.111***
			(0.028)	(0.029)	(0.028)
Team size squared			-0.011***	-0.011***	-0.011***
			(0.003)	(0.004)	(0.003)
Project duration			-0.003	-0.003	-0.003
			(0.003)	(0.003)	(0.003)
Private part.			0.937**	0.994**	0.935**
			(0.463)	(0.483)	(0.461)
Cluster			1.510***	1.545***	1.494***
			(0.144)	(0.149)	(0.143)
Scientific field	YES	YES	YES	YES	YES
Year	YES	YES	YES	YES	YES
Program	YES	YES	YES	YES	YES
Observations	8,685	8,685	8,685	8,685	8,685
rho	-0.76	-0.72	-0.43	-0.37	-0.45
Prob chi2	0.000	0.000	0.045	0.16	0.03

Notes: Robust standard-errors in parentheses, clustered at the project level. The control variables Scientific field, Year and Program are included in the model but the coefficients are not reported. Prob chi2 represents the p-value associated with the Wald test of independence; Ho: The selection equation and the structural equation are independent.

Table 1.26 – Factors that influence the probability to submit a project for the non-directed programs (Heckman Probit coefficients reported)

	(1)	(2)	(3)	(4)	(5)
Novelty		-0.002		-0.001	0.006
		(0.005)		(0.005)	(0.005)
Interdisciplinarity	0.311***		0.307***		0.312***
	(0.033)		(0.033)		(0.034)
Individual variables					
Age	0.071***	0.072***	0.071***	0.073***	0.071***
	(0.006)	(0.006)	(0.007)	(0.006)	(0.007)
Age squared	-0.001***	-0.001***	-0.001***	-0.001***	-0.001***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Male	0.065***	0.071***	0.064***	0.071***	0.064***
	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)
Status					
Professor	0.405***	0.413***	0.402***	0.410***	0.402***
	(0.019)	(0.019)	(0.019)	(0.019)	(0.019)
Associate Researcher	0.427***	0.427***	0.406***	0.407***	0.406***
	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)
Researcher	0.706***	0.709***	0.686***	0.688***	0.685***
	(0.025)	(0.025)	(0.025)	(0.025)	(0.025)
Citations	0.008***	0.008***	0.008***	0.008***	0.007***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Application experience					
nb projects	0.132***	0.136***	0.130***	0.134***	0.130***
	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)
Ongoing grant	-0.325***	-0.327***	-0.329***	-0.331***	-0.329***
	(0.029)	(0.029)	(0.029)	(0.029)	(0.029)
Refusal	0.734***	0.741***	0.735***	0.743***	0.735***
	(0.022)	(0.022)	(0.022)	(0.022)	(0.022)
Lab variables					
Lab size			-0.000	-0.000*	-0.000
			(0.000)	(0.000)	(0.000)
Nb prev project lab	0.003***	0.003***	0.003***	0.003***	0.003***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Rank					
A+			0.064***	0.060***	0.064***
			(0.014)	(0.014)	(0.015)
B			-0.140***	-0.146***	-0.140***
			(0.025)	(0.025)	(0.025)
C			-0.518***	-0.528***	-0.519***
			(0.158)	(0.157)	(0.158)
Missing			-0.033	-0.029	-0.032
			(0.065)	(0.065)	(0.065)
Scientific field	YES	YES	YES	YES	YES
Year	YES	YES	YES	YES	YES
Observations	138,196	138,196	138,196	138,196	138,196

Notes: Robust standard-errors in parentheses, clustered at the individual or project level. The control variables Scientific field and Year are included in the model but the coefficients are not reported. The size of the sample is 138,196 since when a researcher apply to a non-directed program a given year, we deleted the other application to directed programs this same year if necessary.

Table 1.27 – Factors that influence the probability to receive a grant for the non-directed programs (Heckman Probit coefficients reported)

	(1)	(2)	(3)	(4)	(5)
Novelty		-0.014		-0.017	-0.020*
		(0.011)		(0.013)	(0.012)
Interdisciplinarity	-0.294***		-0.269***		-0.281***
	(0.070)		(0.097)		(0.093)
Individual variables					
Age	-0.056***	-0.054***	-0.071***	-0.067**	-0.072***
	(0.017)	(0.019)	(0.022)	(0.028)	(0.021)
Age squared	0.001***	0.001***	0.001***	0.001*	0.001***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Male	0.003	0.004	0.028	0.034	0.027
	(0.036)	(0.038)	(0.046)	(0.053)	(0.045)
Status					
Professor	-0.083	-0.052	0.035	0.090	0.024
	(0.116)	(0.126)	(0.165)	(0.211)	(0.158)
Associate Researcher	-0.108	-0.073	0.017	0.072	0.005
	(0.115)	(0.126)	(0.162)	(0.206)	(0.155)
Researcher	-0.166	-0.107	0.030	0.123	0.010
	(0.185)	(0.202)	(0.262)	(0.335)	(0.250)
Citations	0.002	0.002	0.004	0.004	0.004
	(0.002)	(0.002)	(0.003)	(0.004)	(0.003)
Application experience					
nb projects	-0.060**	-0.055*	-0.041	-0.030	-0.042
	(0.028)	(0.031)	(0.039)	(0.051)	(0.038)
Ongoing grant	0.089	0.066	0.008	-0.030	0.017
	(0.088)	(0.094)	(0.123)	(0.152)	(0.119)
Refusal	-0.409***	-0.374**	-0.279	-0.207	-0.296
	(0.131)	(0.149)	(0.211)	(0.288)	(0.198)
Lab variables					
Lab size			0.000	0.000	0.000
			(0.000)	(0.000)	(0.000)
Rank					
A+			0.044	0.057	0.039
			(0.044)	(0.049)	(0.042)
B			-0.124	-0.146	-0.118
			(0.105)	(0.123)	(0.102)
C			-0.054	-0.110	-0.030
			(0.538)	(0.597)	(0.525)
Missing			-0.153	-0.180	-0.155
			(0.162)	(0.172)	(0.159)
Current application					
Team size			0.242***	0.256***	0.239***
			(0.077)	(0.079)	(0.076)
Team size squared			-0.032**	-0.034**	-0.031**
			(0.013)	(0.013)	(0.013)
Project duration			0.003	0.003	0.003
			(0.003)	(0.003)	(0.003)
Private part.			1.395***	1.491***	1.379***
			(0.481)	(0.504)	(0.477)
Cluster			2.814***	2.965***	2.774***
			(0.534)	(0.585)	(0.523)
Scientific field	YES	YES	YES	YES	YES
Year	YES	YES	YES	YES	YES
Observations	10,376	10,376	10,376	10,376	10,376
rho	-0.75	-0.69	-0.54	-0.45	-0.57
Prob chi2	0.01	0.03	0.18	0.41	0.14

Notes: Robust standard-errors in parentheses, clustered at the project level. The control variables Scientific field and Year are included in the model but the coefficients are not reported. Prob chi2 represents the p-value associated with the Wald test of independence; Ho: The selection equation and the structural equation are independent.

Table 1.28 – Factors that influence the probability to submit a project for the directed programs (Heckman Probit coefficients reported)

	(1)	(2)	(3)	(4)	(5)
Novelty		0.030***		0.030***	0.044***
		(0.006)		(0.006)	(0.006)
Interdisciplinarity	0.559***		0.551***		0.587***
	(0.033)		(0.033)		(0.033)
Individual variables					
Age	0.144***	0.146***	0.146***	0.147***	0.146***
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
Age squared	-0.002***	-0.002***	-0.002***	-0.002***	-0.002***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Male	0.061***	0.070***	0.063***	0.072***	0.064***
	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)
Status					
Professor	0.460***	0.475***	0.459***	0.474***	0.459***
	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)
Associate Researcher	0.319***	0.325***	0.311***	0.317***	0.312***
	(0.019)	(0.019)	(0.019)	(0.019)	(0.019)
Researcher	0.686***	0.692***	0.678***	0.683***	0.674***
	(0.024)	(0.024)	(0.024)	(0.024)	(0.024)
Citations	0.005***	0.005***	0.005***	0.005***	0.004***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Application experience					
nb projects	0.187***	0.191***	0.188***	0.191***	0.185***
	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)
Ongoing grant	-0.102***	-0.104***	-0.105***	-0.106***	-0.104***
	(0.028)	(0.028)	(0.028)	(0.028)	(0.028)
Refusal	0.628***	0.639***	0.626***	0.638***	0.624***
	(0.023)	(0.023)	(0.023)	(0.023)	(0.023)
Lab variables					
Lab size			0.000	0.000	0.000
			(0.000)	(0.000)	(0.000)
Nb prev project lab	0.002**	0.002***	-0.000	0.001	-0.000
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Rank					
A+			-0.037**	-0.040***	-0.033**
			(0.015)	(0.015)	(0.015)
B			-0.166***	-0.180***	-0.164***
			(0.024)	(0.023)	(0.024)
C			-0.208*	-0.225**	-0.208*
			(0.107)	(0.105)	(0.107)
Missing			0.094*	0.096*	0.100*
			(0.056)	(0.056)	(0.056)
Scientific field	YES	YES	YES	YES	YES
Year	YES	YES	YES	YES	YES
Observations	138,066	138,066	138,066	138,066	138,066

Notes: Robust standard-errors in parentheses, clustered at the individual or project level. The control variables Scientific field and Year are included in the model but the coefficients are not reported. The size of the sample is 138,066 since when a researcher apply to a directed program a given year, we deleted the other application to non-directed programs this same year if necessary.

Table 1.29 – Factors that influence the probability to receive a grant for the directed programs (Heckman Probit coefficients reported)

	(1)	(2)	(3)	(4)	(5)
Novelty		-0.039***		-0.027**	-0.036***
		(0.011)		(0.013)	(0.013)
Interdisciplinarity	-0.578***		-0.399***		-0.419***
	(0.065)		(0.088)		(0.090)
Individual variables					
Age	-0.125***	-0.119***	-0.085***	-0.080***	-0.084***
	(0.019)	(0.023)	(0.022)	(0.024)	(0.022)
Age squared	0.001***	0.001***	0.001***	0.001***	0.001***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Male	-0.015	-0.015	-0.004	-0.005	-0.004
	(0.027)	(0.030)	(0.032)	(0.033)	(0.032)
Status					
Professor	-0.213***	-0.187*	-0.010	0.003	-0.007
	(0.081)	(0.097)	(0.067)	(0.073)	(0.067)
Associate Researcher	-0.075	-0.049	0.119*	0.130**	0.120**
	(0.075)	(0.086)	(0.061)	(0.064)	(0.061)
Researcher	-0.234*	-0.173	0.112	0.144	0.118
	(0.127)	(0.149)	(0.094)	(0.102)	(0.093)
Citations	0.002	0.002	0.005***	0.005***	0.005***
	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)
Application experience					
nb projects	-0.071**	-0.060*	-0.005	0.000	-0.002
	(0.028)	(0.033)	(0.022)	(0.024)	(0.022)
Ongoing grant	-0.082	-0.093*	-0.180***	-0.181***	-0.181***
	(0.052)	(0.055)	(0.056)	(0.057)	(0.056)
Refusal	-0.453***	-0.423***	-0.183**	-0.162*	-0.179**
	(0.079)	(0.098)	(0.077)	(0.085)	(0.076)
Lab variables					
Lab size			0.000	0.000	0.000
			(0.000)	(0.000)	(0.000)
Rank					
A+			0.145***	0.145***	0.142***
			(0.029)	(0.029)	(0.029)
B			-0.036	-0.036	-0.041
			(0.059)	(0.060)	(0.059)
C			0.164	0.171	0.166
			(0.266)	(0.268)	(0.266)
Missing			0.097	0.104	0.094
			(0.108)	(0.109)	(0.108)
Current application					
Team size			0.010	0.008	0.010
			(0.022)	(0.022)	(0.022)
Team size squared			-0.001	-0.001	-0.001
			(0.002)	(0.002)	(0.002)
Project duration			-0.005*	-0.006*	-0.005*
			(0.003)	(0.003)	(0.003)
Private part.			0.228	0.271	0.234
			(0.254)	(0.263)	(0.255)
Cluster			1.245***	1.261***	1.245***
			(0.073)	(0.072)	(0.073)
Scientific field	YES	YES	YES	YES	YES
Year	YES	YES	YES	YES	YES
Program	YES	YES	YES	YES	YES
Observations	10,687	10,687	10,687	10,687	10,687
rho	-0.72	-0.64	-0.31	-0.26	-0.309
Prob chi2	0.003	0.02	0.02	0.076	0.02

Notes: Robust standard-errors in parentheses, clustered at the project level. The control variables Scientific field, Year and Program are included in the model but the coefficients are not reported. Prob chi2 represents the p-value associated with the Wald test of independence; Ho: The selection equation and the structural equation are independent.

Table 1.30 – Factors that influence the probability to receive a grant for the sample of teams (Probit coefficients reported)

	(1)	(2)	(3)	(4)	(5)
Novelty team		-0.122***	-0.117***		-0.099***
		(0.036)	(0.036)		(0.037)
Interdisc. team	-0.331**		-0.303**	-0.269*	
	(0.138)		(0.138)	(0.142)	
Members diversity					
Field diversity	-0.239***	-0.235***	-0.214***	-0.247***	-0.245***
	(0.080)	(0.080)	(0.080)	(0.083)	(0.083)
Lab diversity	-0.252***	-0.262***	-0.257***	-0.250**	-0.259***
	(0.095)	(0.095)	(0.095)	(0.099)	(0.099)
Lab quality diversity	-0.184**	-0.183**	-0.178**	-0.213**	-0.211**
	(0.090)	(0.090)	(0.091)	(0.094)	(0.094)
Status diversity	-0.052	-0.047	-0.041	-0.084	-0.080
	(0.088)	(0.088)	(0.088)	(0.091)	(0.091)
Team variables					
Age team	-0.113***	-0.109**	-0.104**	-0.109**	-0.106**
	(0.043)	(0.043)	(0.043)	(0.044)	(0.044)
Age team squared	0.001**	0.001**	0.001**	0.001**	0.001**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
PI male	0.032	0.035	0.033	-0.005	-0.002
	(0.047)	(0.047)	(0.047)	(0.049)	(0.049)
Citations team	0.097***	0.099***	0.107***	0.098***	0.099***
	(0.022)	(0.022)	(0.023)	(0.023)	(0.024)
Top lab	0.206***	0.209***	0.200***	0.203***	0.205***
	(0.045)	(0.045)	(0.045)	(0.047)	(0.046)
Nb grants team	0.055***	0.058***	0.058***	0.055***	0.057***
	(0.019)	(0.020)	(0.020)	(0.020)	(0.020)
Team size	0.036	0.039	0.038	0.019	0.022
	(0.033)	(0.033)	(0.033)	(0.034)	(0.034)
Team size squared	-0.003	-0.003	-0.003	-0.002	-0.002
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Project duration	-0.002	-0.003	-0.003	-0.002	-0.002
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Private part.	-0.041	-0.051	-0.050	-0.034	-0.039
	(0.062)	(0.062)	(0.062)	(0.075)	(0.074)
Cluster	1.523***	1.520***	1.517***	1.486***	1.487***
	(0.083)	(0.083)	(0.083)	(0.086)	(0.086)
Experimental develop.				-0.131	-0.136
				(0.149)	(0.149)
Industrial research				-0.090	-0.098
				(0.088)	(0.088)
Programs	YES	YES	YES	YES	YES
Year	YES	YES	YES	YES	YES
Observations	5,057	5,057	5,057	4,670	4,670

Notes: Robust standard-errors in parentheses. The control variables Programs and Year are included in the model but the coefficients are not reported. The type of research of the project is only known for 4,670 projects.

Chapter 2

The Design And The Impact Of Project Funding In Science: Lessons From The ANR Experience[‡]

[‡]This chapter is based on a paper co-authored with Nicolas Carayol.

Introduction

Governments financially support research carried out in universities and research organizations via various mechanisms. The competitive funding of research projects has been particularly developed in the US since World War II via federal agencies such as the NIH or the NSF.[1] Presumed advantages of that funding mechanism has led countries and institutions around the globe to develop similar policies. Despite the huge amount of public money at stake worldwide and although the way money reaches research presumably affects efficiency, there is still little large-scale systematic evidence about the impact of such fund allocation schemes. Further, competitive allocation of funds to research proposals actually reveals a significant variability whereas the precise rules and goals of the programs are also likely to affect the outcomes. It is thus important to understand how and why the returns may vary with respect to the specific designs of the funding programs.

This chapter provides new clues on these issues, relying on the recent French experience. France is the fifth largest scientific nation worldwide in terms of citations. In 2005, the French government created a dedicated agency, the *Agence Nationale de la Recherche* (ANR) to implement project-based research funding in the country. Our study focuses on the first five years of the ANR's existence (2005-2009) since sufficient post-funding time has now elapsed for some of the first consequences of this policy to be observed. Over this period, this institution received about seventy thousand applications and allocated nearly two and a half billion euros to research projects, the total cost of which amounts to approximately ten billion euros. The ANR was set up as a nationwide generalist player welcoming applications from all disciplines. As alternative sources of funding for professors and researchers' projects were rather limited over that period, this experience offers an excellent opportunity for appreciating the impact of fund allocation on a large scale. It further offers interesting forms of variation as regards funding programs, recipients' characteristics, disciplines, etc. that can allow us to appreciate the differential impact of research project funding for precise differences in their design.

The ANR runs two distinct types of funding programs: directed and non-directed programs. Non-directed programs are more neutral as they welcome applications from all fields of science which are examined by single discipline panels. Directed programs target emerging and promising research areas and/or fields that are suspected to have large potential for future applications. Their calls for proposals are designed by panels mixing top-level representatives of large research institutes and R&D performing corporations with well established scientists and the selection of awardees among applicants is made by ad-hoc interdisciplinary panels. The underlying rationale of directed programs is that the traditional academic incentives for investigating new or interdisciplinary research areas are not strong enough. It is often argued that risk taking, novelty and interdisci-

[1] According to the Board (2016), yearly extramural federal funding of US universities and colleges has exceeded forty billion dollars since 2010. See Stephan (2012) for a detailed overview.

plinarity are under-rewarded because the peer review system would be mainly organized within disciplines and negatively biased toward truly transformative ideas (Braben 2004, Chubin et al. 1990, Wessely 1998). Professors and researchers who respond to incentives (Dasgupta & David 1994, Stephan 2012) and who are autonomous in the choice of their research agendas (Carayol & Dalle 2007, Aghion et al. 2008) may overly refrain from addressing such problems. However nothing guarantees non-neutral funding schemes are more efficient. Applicants to directed programs who are already investigating the targeted fields essentially face more limited competition whereas others may make socially inefficient efforts to comply to the specifics of the calls. Further, directed programs are much more complicated to set up and to be efficiently run. Last but not least, the targeted fields may not have larger potential than totally unanticipated new avenues proposed by applicants. As most research funding programs balance between directed and non-directed rationales, we aim to compare the impact of the two types of programs to appreciate which one of the more neutral (non-directed) or the more interventionist (directed) programs is more efficient.

The main methodological issue of estimating the impact of funding on observational data is to disentangle the selection and the funding effect. Indeed, why university professors and researchers apply to the funding agencies, and why evaluators and committees select them, are also often the same reasons why they are likely to be more productive. Confounding factors are thus likely to affect both funding and the scientific outcomes, which would skew estimates in a naive approach. Jacob & Lefgren (2011) use the grades produced in the evaluation process of NIH grants to account for the selection effect in an IV approach. Those grades given by referees and panel members are intended to capture the variation in projects quality which is uncorrelated with the observables.[2] By doing so, the authors aim to measure the impact of funding, holding constant project quality. Though project quality is brought in by applicants, not by the funding agency, this approach captures the average treatment effect only if unfunded applicants can run projects of similar quality than the ones they submitted. Jacob & Lefgren (2011) indeed argue US biomedical sciences are characterized by a variety of available sources of funding. However, unsuccessful applicants often can not undertake the submitted (and evaluated) projects when alternate funding sources are not available. In those circumstances the extra outcome due to the quality differential between the submitted project and the project undertaken when not selected, is obtained thanks to the funding. As such it should be accounted for in the impact analysis.

As our evidence is characterized by a limited availability of alternative funding sources for projects, the present study thus adopts the conditional difference-in-differences model

[2] Li & Agha (2015) find that the grades significantly explain the scientific performance of the recipients of NIH funds, even controlling for the observables (in particular previous performance). Fang et al. (2016) however have reversed conclusions on the same data set when excluding the projects which got the lowest rates.

developed by Abadie (2005) that allows us to control for both the time-invariant individual fixed effects and the selection on observables.[3] In this model, the identification of the impact of the policy relies significantly on the quality of the observables prior to treatment, on which the fund allocation process is modeled. Fortunately, we were able to assemble detailed information that covers scientists' age, their institutions, fine-grained research fields, and multidimensional publication profiles, which we can use to model fund allocation. Moreover, information is available for almost the whole of the reference population (not just for applicants) as we match the list of applicants with the list of all professors and researchers associated with a laboratory accredited by the Ministry of Research and Higher Education in France. That represents more than thirty thousand tenured scientists while we restrict ourselves to those fields which are sufficiently covered by the publication database we use (Web of Science).[4]

We can, therefore, estimate the impact of receiving an ANR grant using control groups picked either among unsuccessful applicants to the same program and year, or from the whole reference population. There are good reasons to select controls in each way. On the one hand, all applicants self-select and are thus more "similar". On the other hand, picking individuals in the much larger reference population increases the chance of finding controls that are more similar to the treated in terms of the observables (especially as regards publication profiles and detailed scientific fields). In fact, we consider many ways of defining the selection phase which have advantages and drawbacks. Differences lie either in the chosen list of explaining variables (inclusion of individual, laboratory, or trend variables) or in the exclusion rules (picking controls exclusively in the same program, year, or field, or not). We do not postulate that one design of control groups is preferable to the others but test a number of specifications on a placebo parallel path tests before treatment. The best design of the selection stage only considers applicants as potential controls. The selection turns out to be completely unrelated to pre-treatement trends. Balance diagnosis tests show that those properly weighted controls have very similar observables than the treated when they have similar propensity scores. Productivity divergence between controls and treated only starts two years after funding (most project have a three-year period). These remarks converge convincing us that the chosen controls differentiate from the treated for some reasons that are unrelated to their expected scientific productivity, in the absence of treatment. Those controls and scores are thus used to calculate our reference estimations of the impact of funding, but it turns out that using other variants does not qualitatively affect the results obtained.

Overall, we estimate a 15% impact of ANR funding on citations. This is significantly larger than what Jacob & Lefgren (2011) found for the impact of NIH grants (7% impact

[3]The previous literature using the IV (or selection approach) also include Carter et al. (1987), Arora et al. (2000), Arora & Gambardella (2005), Benavente et al. (2012) and Gush et al. (2018). The literature using propensity scores include Chudnovsky et al. (2008) and Azoulay et al. (2011).

[4]Mainly hard and bio-medical sciences but not exclusively as it also includes some social sciences.

on citations) while the average ANR grant is .14 million euros for a total cost of .56 million euros, to be compared with the average NIH grant which is equal to 1.7 million dollars. This difference is likely to be explained by a lower displacement effect due to fewer alternative sources of funding. We attribute to the policy the benefit of recipients working on better quality projects. This is consistent in a context in which alternative sources of funding are scarce. As this value also rests upon funding agencies recruiting good panel members and managing their work efficiently, it is also an outcome of the funding programs that we would like to account for as a component of the policy.

To our knowledge, Azoulay et al. (2011) is the only study comparing the impact of different research funding programs. Their focus is different from ours as they are interested in identifying the differential impact of funding persons vs. projects. They compare the impact of a funding program (HHMI grants) which targets young and promising scholars in the medical fields, with the one of NIH early stage career prizes. They do not use information on the applicants to both funding programs arguing that the recipients of NIH early stage career prizes may, in principle, have applied to the HHMI (same age, country and field). Our data are more complete as we do have information on both the awarded and the unsuccessful applicants for the two programs we compare (directed and undirected). However, the applicants to the two programs are not necessarily the same because directed programs raise barriers to entry. We thus adopt a different estimation strategy than Azoulay et al. (2011) to compare the two types of programs. We use a conditional triple difference approach which literally compares the impacts of the two programs which are themselves estimated as double differences.

We find that the impact of directed programs is rather small (about 6% on citations), while the surplus of impact gained by switching to a non-directed program equals 20%. Non directed programs are significantly more efficient. These programs seem to be able to attract and to pick high quality projects. Though we can not exclude that the directed programs may have delayed impact that we can not fully observe, there is no reason to believe that they achieve their specific goals, with the exception that they do attract and fund professors and researchers who write more novel research articles. The same approach is used to compare other dimensions of programs design, such as the age of the applicants, and find that the impact on younger recipients is significantly larger.

The remainder of the chapter is organized as follows. The data are presented in the second section. Methodology comes next. In the fourth section, we present the selection of controls and the calculation of propensity scores. The fifth section presents our results on the quantification of the average impact of funding. The results concerning the design of funding programs come in the sixth and seventh sections. The last section wraps up and discusses the main results.

I The Data

Data collection

Data collection starts with a list of all researchers and professors associated with one laboratory accredited by the French Ministry of Higher Education and Research around the year 2010, which contains information on 49,225 persons.[5] All of these persons are tenured, whenever as full or assistant professors, assistant researchers or research directors. Once all individuals for which we do not have full and consistent information (status, institutional employer, laboratory, age, etc.) have been excluded, we are left with 48,328 persons. This list has then been matched to the names of the authors of scientific articles, letters and reviews (on the basis of their surname and first name initials) in the Thomson-Reuters ISI Web of Science, a well-known database which gathers all the documents published in the main scientific journals. The publication period covered in this study goes up to and including year 2012. Thus, the last publication year considered (2012) stands three years after the last funding year (2009) and seven years after the first funding year (2005). We collected more than nine million distinct authorships (listed author×document) which received more than forty million citations. As these large publication records show, we are faced with a huge homonymy problem due to the absence of any reliably unique identifier of researchers in publication databases. A disambiguation algorithm has thus been developed based on a "seed + expand" methodology (Reijnhoudt et al. 2014). Basically, this algorithm works as follows: in a first step (seed), the algorithm validates articles by imposing strong conditions, particularly on the field and institutions, which should be consistent with what we know for each person. At this stage, the goal is to minimize false positives. In the second step (expand), the algorithm uses the information on the articles already validated in the seed step, to accept other articles which did not fully meet the conditions of the seed step. Typically the information used concerns the co-authors, the references and the keywords. New papers are validated either because, besides some of the first-step conditions which are maintained, they have the same co-authors or cite the same references as already validated articles. The program then iterates up to some point. In order to evaluate the quality of this disambiguation process, we have constituted a benchmark of nearly 300 scientists who have created an ORCID number and are thus likely to have disambiguated their own publications. Detailed information on the algorithm and on the quality of the disambiguation are presented in the Appendix G. By the end of the disambiguation process, 1.2 million author×documents have been validated (733 thousand distinct articles), that is approximately 13% of the initial set.

[5] Are thus excluded all the tenured researchers and professors who are not associated to a lab, and those associated to laboratories in schools funded only by other ministries (such as the ministries of industry, agriculture or defense), or to laboratories solely associated to national research institutes (such as CNRS or CEA internal labs).

The affiliation of professors and researchers to scientific fields of investigation is based on a fine-grained organization of science in France into peer groups called "sections". Such sections are specific to the institutional employer, either a national research institute (such as CNRS or INSERM) or the Ministry of Higher Education and Research for all professors employed in universities and schools. Each section members elect a national committee which usually accredits PhDs for recruitment (or sometimes even recruits directly), evaluates individuals, allocates promotions, etc. Most of the time, sections tend to be organized around specific disciplinary orientations.[6] We computed, for each section, the percentage of professors and researchers for whom no article was found in the database. On the basis of this information, we excluded a long list of sections, mostly in the fields of humanities and social sciences. We suspect that these disciplines are not well covered by the database, either because scientific journal articles are not the main outcomes of their research, or because the principal journals of these disciplines are not well covered by the database. This leaves us with 31,081 persons.

The ANR provided us the list of all applications from 2005 to 2009, comprising 67,812 partners $\times$ applications. A project "partner" is defined as an institution which will directly receive the planned funds from the ANR if the application is successful. Each partner has its own scientific coordinator. Multipartner projects have only one project coordinating partner, whose scientific coordinator is the project PI. In multipartner projects, each partner receives its funds directly from the ANR. Each partner coordinator is fully responsible for the engagement of the funds received by his/her institution and thus enjoys significant autonomy. Keeping only the partners $\times$ applications emanating from academia and for which the variables of interest are correctly documented (scientific coordinator's surname and first name, the partner, funding decision, amount, and duration), leaves 54,852 partners $\times$ applications. The success rate is 30%. The total amount allocated is 2.4 billion euros, but the expected total cost of the funded projects is 9.5 billion euros because the ANR funds only the marginal cost of the projects it supports for public partners.[7] The median fund per partner is 136,000 euros, while the mean is 138,000 euros. The mean total cost per partner is 545,000 euros.

We next basically matched the list of scientific coordinators of all ANR applications with the personnel list obtained previously. Two types of matching were performed subsequently: an exact matching and a fuzzy one.[8] In the event of homonymy in the full

[6]For a few specialized research institutes, the specialty of the sections is not straightforward, and we had to develop specific strategies. For instance, for INRA (the national research body dedicated to agricultural research), the allocation to disciplines has been performed on an individual basis.

[7]The grants cover the wages of the non-tenured personnel hired for the purpose of the project and overheads limited to 4% of the grant. The total costs typically include the grant and all the resources included in the project, in particular the salaries of the tenured researchers and professors paid by the research institutes and universities.

[8]Fuzzy matching authorizes small variations in the surnames and first names and then requires manual verification and cleaning, basically comparing individual information and project information before validation.

initial list of scientists, a manual check was made, based on the consistency between the discipline of the scientist and the project description, and between the employer of the scientist and the project partner. This matching allowed us to find, in the list of the 31,081 professors and researchers, the scientific coordinators of 46.2% of all applications, 45.5% of the funds, and 46.9% of the total amount of money allocated.

It turns out that more than one third (10,722) of all these persons applied as scientific coordinators of the partners involved in the projects submitted between 2005 and 2009, and that 18.6% (5,831 persons) obtained at least one grant (4,892 applicants were never funded). Therefore, two third (20,498 persons) did not apply. The age distribution of the three populations (reference population, applicants and funded) is similar, though the 35-50 years-old (in year 2007) are proportionately more numerous among the applicants and the funded[9]. Researchers and full professors are more likely to have applied at least once. Researchers from CNRS and INSERM apply more often and their applications are more likely to be successful. The applicants identified have applied on average 2.4 times over the period (25,364 applications). The distribution of applications is asymmetric, with most professors and researchers not applying or applying only once, while some apply many times. On average, the applicants obtained 1.2 grants over the five years considered (12,757 funds allocated). Like the applications, the funds are also unevenly allocated across the population: More than 75% of the applicants received only one funding, while a few got many. In this study, we will consider only the first funding for those who got multiple grants. There are two types of programs: directed programs that have a specific directed orientation, and non-directed ones which are fully open to any application. While half of the applications go to directed programs and the other half to non-directed programs, directed programs account for 65% of the grants allocated, because these programs have significantly higher rates of success.

When we break down applications by discipline, we observe that the highest rate of application is found for physics (with more than one application per scientist), followed by fundamental biology (.94), chemistry (.91) and applied biology and ecology (.90). The lowest rate is found for mathematicians who applied only one-third time on average. The highest average rate of funding can be observed for applied biology and ecology (with .34 funds per scientist). Physics follows immediately (with .32 funds per capita). These two fields differ strongly, however, in terms of supporting programs: physics is most often funded by non-directed programs, whereas nearly two-thirds of the funds allocated to applied biology and ecology come through directed programs. Similarly, fundamental biology, medicine and engineering sciences are mostly funded by directed programs, while the sciences of the universe and mathematics are most often funded by non-directed programs.

[9]Some descriptive tables and figures are presented in the Appendix A

Outcome variables

ANR funding is intended to sustain the scientific production and excellence of the awardees. Though this can not be observed directly, different measurements of scientific outcomes can assess its most important dimensions.[10] We build three variables that are labeled respectively Volume, Citations and Impact Factor. Though they are not independent, these indicators are distinct and proxy different dimensions of scientific production and excellence. Volume sums the number of articles published, each being adjusted by the number of co-authors (fractional counts). It relates more to the volume/quantity of scientific production. Impact Factor weights each article by the average number of citations which papers published in the journal that year received on average (again in a three-year window).[11] That variable captures the capacity to publish in well-established journals. Citations weights each article by the number of citations it received (in a three-year time window). As such this indicator captures the impact of each article on the scientific literature and thus corresponds more directly to scientific excellence. All three indicators may be significantly affected by field differences for a number of reasons, but, as these differences are time invariant, they are controlled for in the difference-in-differences design.

Publication data also prove to be very helpful in investigating the collaboration behaviors of professors and researchers (Wuchty et al. 2007). We use the number of authors of the article to evidence the size of the research teams, information which is averaged for each given period and person to obtain variable Average Team Size. Collecting all collaborators' names and initials over given time periods and dropping double counts, we also compute the total number of distinct co-authors, labeled Coauthors. This number proxies the size of the collaboration network. We also compare the sets of collaborators between two consecutive time periods to assess the number of new co-authors they are working with, labeled New Coauthors. The addresses of the authors' institutions can be used to assess the capacity of professors and researchers to extend their collaboration networks at the international level. The variable International Collaborations equals the number of articles that have at least one foreign address.

Descriptive statistics

Some descriptive statistics of the outcome variables measured on the whole data set are presented in Table 2.1. Mean outcomes are presented separately for the three years before and three years after the year of reference and for the three defined groups, the non-applicants, the not granted applicants and the granted ones. The reference year is the

[10] Details on the calculation of all outcome variables are presented in Appendix B.

[11] This weighting scheme is very close to, but distinct from, the traditional Journal Impact Factor which divides the number of citations received in a given year (thus to articles published that year but also to those published previously) by the number of articles published that year. Therefore our approach is less sensitive to the yearly variations in the average quality or in the number of articles published.

year of first successful application for the funded. There are exactly 5,831 distinct persons in our data set that have been funded by the ANR. We do not consider their subsequent successful applications, nor their unsuccessful ones. There are 9,652 distinct applications of the 4,892 persons who applied at least once, but were never funded. We consider all the applications of the never funded applicants, and for each of those applications, the before and after periods are defined according to the application year. As there is no specific reference year for the 20,498 non-applicants, they are considered for each of the five years of the study, and the before and after periods are defined accordingly. Incomplete information about the identity of co-authors before year 2002 prevents us to compute the New Coauthors variable in the period preceding the year of reference.

The means measured in the three-year period after the year of reference are always higher than those in the previous period, but differences in magnitude are observed according to groups. The subset of non-applicants tends to publish more articles after the year of reference (1.64 against 1.42 articles in the previous period, in fractional counts), of a higher accumulated journal impact factor (2.98 against a mean impact factor of 2.43 before) and which receive more citations (6.19 against 5.74 citations before). This positive evolution is also observed for not granted applicants (2.4 against 2.12 articles in the past period), who publish more articles than the non-applicants but less than the granted ones (3.16 against 2.8 articles in the past period). Their publications also received more citations (10.03 against 9.47 citations before) and are associated to higher journal impact factors (4.84 against 3.97 previously), but again in a lesser extent than granted applicants (16.19 against 14.45 citations in the past period and a mean impact factor of 7.14 against 5.77 before).

The collaboration profiles also contrast between groups. The non-applicants are characterized by a larger average team size than the other groups in the past period (10.65 co-authors compared with 7.34 co-authors for the not granted applicants and 7.55 co-authors for the granted applicants). This difference is sharpened in the subsequent period, with an average team size of 20 co-authors for the non-applicants compared to only around 11 co-authors for the other groups. Granted applicants seem however to collaborate more often with different authors, as indicated by the annual number of co-authors. In the subsequent period, they collaborate on average with 81 different co-authors (40.10 individuals in the last period) compared to 65.77 and 57.47 co-authors for the not granted applicants and the non-applicants respectively (34.17 and 31.41 co-authors respectively in the last period). Those numbers of co-authors may look large. They are however consistent with the average team sizes (note the average number of co-authors per paper are always above ten in the “after” period). Moreover, the averages are driven by outliers in those disciplines characterized by very large author teams (the median numbers of co-authors are significantly lower than the means). Granted applicants tend to collaborate more with researchers from abroad (8.22 times on average in the subsequent period against 6.64 in the last period), compared to not granted applicants (6.62 times against 5.34 in the last period) and non-applicants (5.93 times against 4.71 in the last period). In the subsequent period, granted applicants also collaborate more often with partner they never worked with before, with an average of 18.20 new co-authors, compared to 16.05 new co-authors for not granted applicants and only 12.45 new co-authors for the non-applicants. Finally, we observe only little differences between groups in the propensity to address new problems.

II Identifying the impact of funds: Methodology

Controlling selection on observables

In this paper, as we focus on the effect of receiving an ANR award on successful applicants, we are interested in the so-called average treatment effect on the treated individuals, which is defined as follows:

$$ATT = E\left(Y(1) - Y(0)|\, T = 1\right), \tag{2.1}$$

where $Y(1)$ denotes the production when the applicant is funded, while $Y(0)$ refers to the counterfactual, i.e. the production if the applicant had not been funded. The event noted $T = 1$ means treatment occurs. The problem is that the counterfactual outcome is non-observable: either he/she is funded, or is not, but not both.

Propensity scores can help reduce the bias related to the selection on observable characteristics. Rosenbaum & Rubin (1985) show that, under the ignorability condition which states that adjusting for a set of covariates X is sufficient to remove all confounding factors, controlling for the propensity scores is sufficient. The propensity score $P(X)$ is defined as the probability of being "treated" (obtaining a grant in our case) given X: $P(X) = P(T = 1|X)$, with $0 < P(T = 1|X) < 1$. The propensity scores are reliably estimated when the conditional independence assumption (CIA) is verified.[12] It states that the potential outcome is independent of the treatment status, conditional on the propensity score. In other words, the treated individuals would have reached the same outcome levels as the controls having the same propensity score, if they had not been assigned to the treatment:

$$E\left(Y(0)|\, T = 1,\, p(X)\right) = E\left(Y(0)|\, T = 0,\, p(X)\right) = E\left(Y(0)\right). \tag{2.2}$$

This equation can be rewritten as:

$$Y(0) \perp T|p(X). \tag{2.3}$$

[12] It is also known as Weak Unconfoundedness for the ATT.

Such an assumption, which cannot be tested directly, implies that there is no confounder influencing both the assignment of the treatment and the outcome that is not included in X. Heckman et al. (1997) show that the non-inclusion of a relevant covariate causes the introduction of a bias in the estimated impact. In other words, the CIA assumption is valid only if all the covariates which influence both the treatment and the outcome variables are included in the set of explanatory covariates used for the estimation of the propensity scores.

Therefore, the covariates that are included in the vector X, which is used for estimating the propensity scores, need to be selected with caution. In this study, we use an "agnostic" approach whereby we investigate several specifications of the selection model that we test later.

Matching and weighting

Different methods using the estimated propensity scores can be applied to remove the bias due to the differences between the observed characteristics of the treated and those of the untreated individuals. In this paper, we consider two matching procedures, nearest neighbors matching with replacement, and kernel matching, as well as inverse probability of treatment weighting (IPTW).

In the nearest neighbors matching, each treated individual is assigned its most similar controls (up to five) in terms of propensity score. To improve the quality of the matching, a caliper width is specified, which restricts the selection of the controls within a caliper around the propensity score of the treated individual (to avoid capturing controls that are too distant). The caliper value is calculated in line with Cochran & Rubin (1973), who tested the bias reduction when applying a caliper width $c = a\sqrt{(\sigma_1^2 + \sigma_2^2)/2}$, along with σ_1 and σ_2, which are the standard deviations of the propensity scores among the treated individuals and the controls respectively, as well as with a as a positive parameter. Following Rosenbaum & Rubin (1985), we set $a = 0.2$ which is expected to remove around 99% of the bias. Using a caliper condition however reduces the subset of available controls. Note that treated individuals will be excluded from the analysis if no control meets the imposed conditions (the caliper or the common support restriction).

Unlike the nearest neighbors approach which assigns the same weight to all controls of a given treated individual, the kernel matching approach assigns a different weight to each control, which is inversely proportional to the difference between its propensity score and that of the treated individuals. The kernel method provides an interesting solution when the nearest controls have very different propensity scores to those of the treated individuals. Frölich (2004) argues that kernel matching is always preferable to nearest neighbors matching. We exclude observations with extreme propensity score values. Following Imbens & Wooldridge (2009), we remove all individuals i, such that $p(x_i) > .9$ or

$p(x_i) < .1$. We also apply the common support restriction, which implies that we do not consider controls with a lower propensity score than the lowest score among the treated individuals (Dehejia & Wahba 1999).

Robins et al. (2000) and Hirano & Imbens (2001) argue that the controls with higher probabilities of being treated are likely to be under-represented in the control population (because they are likely to have been treated), whereas the controls with lower propensity scores are likely to be over-represented. To correct for this bias, the authors suggest weighting the controls by the inverse of the probability of being treated. The weights allow under-represented controls (because they are likely to have been treated) to have a more important role in the analysis as compared to the controls who have a low probability of receiving the treatment (who are thus likely to be over-represented). Hirano et al. (2003) argue that this approach is more efficient.

Conditional difference-in-differences

So far we have considered that observed heterogeneity was sufficient for explaining the selection into treatment. However, in the applicants' CVs or in their project proposals, the selection committees and the external solicited referees can find relevant information that cannot be observed in our data, but which reveals their ability to perform in science. If this occurs, and if it influences the selection, then propensity scores are not sufficient for identification. However, if these unobserved variables are time-invariant, such as personal fixed-effects, then time differentiation can be used to solve the problem. The relevant approach is the so-called difference-in-differences methodology, which basically compares the variation in the performances of the treated individuals and the controls, before and after treatment. The outcomes variables are calculated by pooling together the information on the three years before and the three years after the year of funding.[13] Therefore, the publication outcomes issued in the year of funding are not considered. The three-year window ensures that the post-funding publication period considered is complete, even for the last funding year considered (2009) because publication data are available until 2012.

In the context of our application, we conceptualize scientific outcome as given by:

$$Y_{i,t} = T_{i,t} \times \delta + \eta_t + \theta_i + \mu_{i,t}, \tag{2.4}$$

where i refers to the professors or researcher, $t = 1, 2$ denotes the time period (pre vs. post treatment), $T_{i,t}$ is the treatment dummy, η_t is a time dummy equal to one in the post treatment period $t = 2$, θ_i is a fixed individual effect, and $\mu_{i,t}$ is the error term. The term

[13] In principle, this analysis could be done on a yearly basis. However, we follow Bertrand et al. (2004) who show that using only two periods is preferable because it reduces serial auto-correlation.

δ is the impact of funding that we aim to estimate. Assuming we remove confoundedness ($cov\left(T_{i,t}, \mu_{i,t}\right) = 0$), δ can be properly estimated in a difference-in-differences approach.

To do so, we use the conditional difference-in-differences model (Abadie 2005) that combines a treatment selection model based on the estimation of the propensity scores with the difference-in-differences method. The estimation of the impact can be calculated as follows:

$$\hat{\delta} = \frac{1}{|N_T|} \sum_{i \in N_T} \omega_i \left(Y_{i,1} - Y_{i,0}\right) - \frac{1}{|N_C|} \sum_{j \in N_{NT}} \omega_j \left(Y_{j,1} - Y_{j,0}\right), \tag{2.5}$$

where N_T denotes the set of treated individuals and N_C the set of controls. $Y_{i,t}$ is the outcome variable observed in period t, with $t = 1$ in the period after the treatment assignment, and $t = 0$ in the period before treatment. The weights ω_j are defined according to the chosen matching or weighting method. When the nearest neighbors or the kernel methods are chosen, the treated individuals have a unitary weight ($\omega_i = 1$) when included and the controls have a total weight which is accumulated over the treated individuals to which they are associated: $\omega_j = \sum_{i \in N_T} \frac{1}{|M(i)|} \omega_{j,i}$, with $\omega_{j,i}$ the weight of control j vis-à-vis treated agent i, and with $M(i)$ the set of controls for treated agent i. With the IPTW approach, the weights are calculated following a slightly different logic as controls are no longer specifically associated with given treated individuals. They are calculated as follows: $\omega_i = T_i + \frac{(1-T_i)p(x_i)}{1-p(x_i)}, \forall i \in N_T \cup N_C$, with $T_i = 1_{\{i \in N_T\}}$ the treatment dummy and $p(x_i)$ the propensity score of agent i (Robins et al. 2000, Hirano & Imbens 2001).

III Selection on observables

In this section we first provide descriptive statistics on the variable used to model the selection stage, before presenting the models and the tests.

Descriptive statistics

Some descriptive statistics of the selection variables are presented in Table 2.2. The three population sets (non-applicants, not-granted applicants and granted applicants) are exposed separately. In Table 2.2, we also distinguish, among the not granted applicants, the ones who applied to directed vs. to non-directed programs. The same is done for the granted. This distinction makes no sense for the non-applicants. Out of the 5,831 professors and researchers who have been funded by the ANR, 3,385 got their first application thanks to a directed program, and 2,446 thanks to the non-directed program. There are a few cases for which a first grant from the directed and from the non-directed programs occur simultaneous. When this happens, one is randomly selected while others are excluded from the sample. We do not consider the subsequent successful applications

Table 2.1 – Descriptive statistics on outcome variables among non-applicants, unsuccessful applicants and granted ones, before and after the reference year.

Variables		Non-Applicants		Not Granted Applicants		Granted Applicants	
	stat.	Before	After	Before	After	Before	After
	mean	1.42	1.64	2.33	2.58	2.80	3.16
Volume	*med.*	0.53	0.68	1.45	1.58	1.83	2.05
	s.d.	(2.66)	(3.01)	(3.04)	(3.40)	(3.35)	(3.72)
	mean	2.43	2.98	4.69	5.34	5.77	7.14
Impact Factor	*med.*	0.59	0.88	2.53	2.89	3.14	3.93
	s.d.	(5.41)	(6.40)	(7.02)	(7.86)	(8.19)	(9.95)
	mean	5.74	6.19	11.14	10.58	14.45	16.19
Citations	*med.*	0.92	1.12	4.90	4.27	6.68	7.13
	s.d.	(15.16)	(15.99)	(20.19)	(20.24)	(23.97)	(27.45)
	mean	10.65	19.93	7.24	10.29	7.55	11.34
Av. team Size	*med.*	5.14	5.60	5.50	6.00	5.29	5.91
	s.d.	(57.67)	(152.21)	(21.80)	(70.70)	(27.47)	(83.95)
	mean	31.41	57.47	37.37	71.32	40.10	81.82
Coauthors	*med.*	13	20	22.00	37.00	24	44
	s.d.	(51.25)	(107.04)	(49.56)	(104.06)	(50.85)	(116.56)
	mean	4.71	5.93	5.92	7.29	6.64	8.22
Internat. Collab	*med.*	1	2	3	3	3	4
	s.d.	(10.65)	(14.77)	(10.49)	(13.35)	(11.04)	(13.94)
	mean		12.45		17.34		18.20
New Coauthors	*med.*		8		15		16
	s.d.		(11.55)		(12.55)		(12.84)
	mean	10.58	10.84	10.78	11.01	10.73	10.99
New problems	*med.*	10.73	10.96	10.85	11.10	10.79	11.07
	s.d.	(0.86)	(0.85)	(0.78)	(0.77)	(0.75)	(0.74)
Nb. of observations		102,490		9,652		5,831	

Notes: The "before" ("after") columns refer to the three years which precede (follow) the year of reference. It is the year of first successful application for the funded. There are exactly 5,831 distinct persons in our data set that have been funded by the ANR. We do not consider their subsequent successful applications, nor their unsuccessful ones. There are 9,652 distinct applications of the 4,892 persons who applied at least once, but were never funded. We thus consider all the applications of the non funded applicants, and for each of those applications, the before and after periods are defined according to the application year. As there is no specific reference year for the 20,498 non-applicants, they are considered for each of the five years of the study (102,490 observations), and the before and after periods are defined accordingly. Incomplete information about the identity of co-authors before year 2002 prevents us to compute the New Coauthors variable in the period preceding the year of reference.

of the granted, nor their unsuccessful ones. The never successful applicants are considered for each of their applications, 4,085 to directed programs, and 5,567 to non-directed ones. The non-applicants are considered five times, that is, once for each of the funding years. As all presented statistics are time-variant, they are calculated for each considered year.

As age is likely to explain both the probability to apply and the probability of being granted, we consider age at the time of application (Age). The different subpopulations however do not differ significantly in terms of their average or median age, but the applicants (granted or not) to directed programs are two-to-three years older on average. The number of articles (fractionned to account for co-authorship) published in the previous three years is intended to capture recent research intensity (Articles). It is significantly larger for the applicants than for the non-applicants. Among applicants, it is 20% larger for the granted than for the not granted. We use the number of citations to those articles (keeping the fractional counting) received in a three-year period after publication to account for the scientific impact of recent research (Citations). Similar differences are found between the three groups (the granted perform better than not granted applicants, who perform better than non-applicants). However, we now also find that applicants (both granted or not-granted) to non-directed programs have larger citations records than their counterparts in the directed programs. The total number of citations received over their career, recorded since 1999, accounts for long-run reputation (Total Citations). It may affect both self-selection (applying or not) and the odds of passing the formal selection process. We observe similar differences as for Citations. Note that if non-applicants have significantly lower scores on average, this is mainly explained by a large proportion of low performing individuals in this population as the median equals 9, which is only 11% to 16% of the median in the different groups of applicants. The largest Impact Factor of the journals which published their three previous years papers accounts for the capacity to publish in well established journals (Max Impact Factor). We observe a neat difference between the non-applicants and the applicants on this variable (three times greater), but differences between sorts of applicants are rather limited.

A series of variables are employed to capture the characteristics of the research environment. The average number of articles per capita in the laboratory accounts for the intensity of research activity in the close professional environment (Av. Lab Articles). To account for the presence of one particularly reputed member of the lab, which could affect the probability to apply and to be selected, we use the maximum number of citations reached among lab members (Max Lab Citations). We find that applicants have more intense research environments than non-applicants and are more likely to have a star in their lab. Larger labs often have larger supporting staff which may affect the probability to apply and the quality of the project proposal. We thus use the number of faculty members to capture the size of the lab (Lab Size). There are however limited differences between the various sorts of applicants in these respects. The average laboratory size is

Table 2.2 – Descriptive statistics on selection variables for non-applicants, not granted applicants and granted ones, by program type (directed or non-directed).

Variables		Non-Applicants	Not Granted Applicants		Granted Applicants	
	stat.		Directed	Non-Directed	Directed	Non-Directed
	mean	42.87	45.30	42.73	44.44	42.50
Age	*med.*	41.00	45.00	41.00	44.00	41.00
	s.d.	(10.25)	(8.09)	(8.55)	(8.16)	(8.35)
	mean	6.49	10.92	10.26	12.21	12.17
Articles	*med.*	2.00	7.00	6.00	8.00	8.00
	s.d.	(12.61)	(14.64)	(13.16)	(14.98)	(14.19)
	mean	37.50	58.25	64.45	75.48	80.94
Citations	*med.*	4.00	23.00	26.00	30.00	37.00
	s.d.	(117.74)	(111.97)	(129.73)	(133.83)	(138.06)
	mean	78.34	132.60	148.28	153.85	164.80
Total Citations	*med.*	9.00	57.00	67.00	63.00	79.00
	s.d.	(231.82)	(253.57)	(272.00)	(276.58)	(271.08)
	mean	2.53	3.88	4.47	4.46	4.95
Max Impact Factor	*med.*	1.26	3.25	3.65	3.36	3.74
	s.d.	(3.42)	(3.71)	(4.11)	(4.57)	(4.53)
	mean	7.68	9.08	9.36	8.97	8.57
Av. Lab Articles	*med.*	6.64	8.38	8.68	8.18	8.14
	s.d.	(5.96)	(6.39)	(5.65)	(5.97)	(4.88)
	mean	369.09	409.39	461.10	408.82	447.05
Max Lab Citations	*med.*	247.00	305.00	358.00	321.00	333.50
	s.d.	(422.83)	(421.27)	(421.91)	(409.88)	(445.70)
	mean	52.01	51.58	51.99	54.12	56.20
Lab Size	*med.*	42.00	41.00	45.00	42.00	48.00
	s.d.	(40.41)	(41.38)	(36.98)	(43.91)	(39.12)
	mean	2007.00	2007.13	2007.95	2006.64	2006.68
Application Year	*med.*	2007.00	2007.00	2009.00	2006.00	2006.00
	s.d.	(1.41)	(1.18)	(1.20)	(1.32)	(1.39)
Nb. of observations		102,490	4,085	5,567	3,385	2,446

Notes: The "Directed" ("Non-Directed") distinction does not make sense for the 20,498 non-applicants. They are considered in this table as five distinct potential controls, one for each of the five years of the study, and the statistics are computed accordingly. Out of the 5,831 who have been funded by the ANR, 3,385 got their first application thanks to a directed program, 2,446 thanks to the non-directed program. We do not consider their subsequent successful applications, nor their unsuccessful ones. There are 9,652 distinct applications of the 4,892 persons who applied at least once, but were never funded. 4,085 applications to directed programs, and 5,567 to non-directed ones.

pretty similar between groups of applicants and with the non-applicants. It is however larger for the applicants (treated or not) to non-directed programs.

Selection models and tests

Eight different designs have been retained after some tests for selecting potential controls and calculating propensity scores. Table 2.3 synthesizes the different models.[14] All logit estimations regress the treatment dummy on individual variables. The latter include age which may affect the odds of being granted as well as the various variables discussed above on individual past publication profiles (number of publications in the last three years, number of citations to articles published over the same period, highest impact factor, and total number of citations received over their career so far). The designs for calculating propensity scores differ, however, in several respects. Publication trend variables in the years preceding treatment are included in some logit regressions so as to capture the recent dynamics of scientific production before treatment. Some designs exclude all non-applicants, while others select controls from within the reference population as a whole. Some, however, require the controls to be in the same section as the treated individuals,[15] while others do not. Let's recall that the section allows to control for both the detailed scientific field and the employer (a specific national research institute or any university). Since the quality of the research environment is one of the selection criteria of the ANR, we have also considered the inclusion of laboratory variables among the regressors,[16] such as the ones presented above: the average research intensity of the lab, the size of the lab, and the presence of an outstanding reputation member. Last but not least, in some designs the directed and non-directed programs are considered jointly while in others, logit regressions are performed by program type, basically assuming that the selection mechanisms of directed and non-directed programs are distinct, based on different weights given to the observables, and even on different observables.

The difference-in-differences identification relies on the parallel path hypothesis, that is, the treated individuals would have had production paths parallel to the ones of their controls if they had not been treated. This hypothesis cannot be tested comparing controls and treated outcomes, since the counterfactual of treated is not available after treatment.

[14]For each design, the three weighting methods (five nearest neighbors, kernel and inverse probability of treatment weighting) have been tested. That makes 24 estimations. Details are presented in Appendix C.

[15]Then propensity scores are computed separately for each section. This implies that controls are exactly in the same field of study than the treated and have broadly the same status (researcher or professor) and the same employer, that is the Ministry of higher education if the treated is professor, or a specific research institute if the treated is a researcher.

[16]A potential issue is that laboratory variables are observed after funding (approximately around year 2010). Therefore, if the consequence of funding is a mobility in a different lab (potentially of a higher quality), the impact of the funding may actually be underestimated (because the treated would be basically compared to controls in higher quality labs). Our results show, however, that the estimations retaining lab variables do not differ significantly and thus that lab variables do not lead to an underestimation.

Table 2.3 – Synthesis of the eight specifications of the propensity score model.

	1	2	3	4	5	6	7	8
Restriction on the controls								
All the reference population	X	X						
Only applicants			X	X	X	X	X	X
Exact matching								
Section (detailed field & employer)	X	X						
Field & research institute			X	X		X	X	
Program type					X			X
Covariates explaining the treatment								
Individual covariates	X	X	X	X	X	X	X	X
Laboratory covariates		X		X			X	
Trend covariates						X	X	X

Notes: The reference selection model is the eighth.

However, the parallel path hypothesis can be tested between different periods before treatment (Imbens 2004, Abadie 2005). Therefore, we estimate a hypothetical impact of the treatment on the treated individuals, between two distinct periods before treatment: $t_0 - 3$ and $t_0 - 1$, where t_0 stands for the year of funding. The goal is to verify that the variations in outcomes of the treated individuals before treatment are not significantly different from those of the controls. Performing such a test on all estimations, we find that the predicted impact of the treatment is always very small and never significant. A similar test is performed between $t_0 - 3$ and t_0. This placebo test is more helpful in sorting out candidate estimations, and thus in selecting our reference propensity score estimations. The propensity scores that lead to the most parallel paths are those for which the controls are applicants exclusively; the estimations are performed distinctly between program types (directed vs. non-directed); the laboratory variables are not included; and production trend variables before treatment are included. The variables used in the two models (for directed and for non-directed programs) are presented in Tables 2.4 and 2.5.

Following the recent literature we will use the IPTW weighting scheme as our reference estimations but consider the other weighting methods. In fact, the results with the nearest-neighbors and the kernel matching methods remain essentially the same as the ones obtained with the IPTW. The results of the parallel path test for our preferred propensity score specification are presented in Table 2.6. Appendix D presents the same parallel path tests for the other specifications.

The conditional independence assumption on which our identification strategy is based implies that the treatment dummy is independent of the variables included in the logit model, conditional on the propensity score. Therefore, if the estimated propensity scores are correct, we expect that controls and treated individuals do not significantly differ with respect to the explanatory variables when they have similar propensity scores. Dehejia

Table 2.4 – List of covariates used for the propensity score estimation in the reference model, for directed programs.

Variable Description
Age at the time of application
Number of publications in the previous 3 years
Number of citations to papers published in the previous 3 years
Maximum Impact Factor in the previous 3 years
Total number of citations to papers published since 1999
The specific directed program
Year of the application
Interaction between the specific directed program and the application year
Variation in absolute terms in the number of publications $(t-3, t)$
Variation in percentage points in the number of citations $(t-3, t-1)$

Notes: All outcome variables (apart from the total number of citations) are adjusted for co-authorship (fractional counting) and categorized in four classes (four dummies are created): top 10%, next 20%, next 30% and last 40% . The distribution of the variables is restricted to researchers of the same scientific field (31 disciplines used).

Table 2.5 – List of covariates used for the propensity score estimation in the reference model, for non-directed programs.

Variable Description
Age at the time of application
Number of publications in the previous 3 years
Number of citations to papers published in the previous 3 years
Maximum Impact Factor in the previous 3 years
Total number of citations to papers published since 1999
Large scientific disciplines dummies
Dummies when an university or a specific research institute is the employer
Variation in absolute terms in the maximum Impact Factor $(t-3, t)$
Variation in percentage points in the maximum Impact Factor $(t-3, t-1)$

Notes: All outcome variables (apart from the total number of citations) are adjusted for co-authorship (fractional counting) and categorized in four classes (four dummies are created): top 10%, next 20%, next 30% and last 40%. The distribution of the variables is restricted to researchers of the same scientific field (31 disciplines used). The large scientific disciplines dummies are Life sciences, Medicine, Chemistry, Physics, Science of the Universe, Engineering, Mathematics, Information science, Human & social sciences. The specific research institutes are: CNRS, INRA, INRIA, IRD, and INSERM.

Table 2.6 – Parallel path test: Difference-in-differences estimates of the mean effect of treatment on various production variables with the reference specification of the selection stage (calculated from $t-3$ to $t-1$ and from $t-3$ to t).

	from $t-3$ to $t-1$			from $t-3$ to t		
	5 nn	kernel	iptw	5 nn	kernel	iptw
Volume	-.00954	-.00809	-.00727	.00078	-.00101	-.00139
	(-1.39)	(-1.27)	(-1.13)	(0.11)	(-0.15)	(-0.21)
Citations	-.00454	-.0012	-.00268	-.01868	-.02018	-.02246
	(-0.3)	(-0.09)	(-0.19)	(-1.23)	(-1.45)	(-1.58)
Impact Factor	-.00515	-.00383	-.00588	-.00626	-.0040	-.00484
	(-0.6)	(-0.47)	(-0.72)	(-0.69)	(-0.47)	(-0.56)

*Notes: Conditional difference-in-differences results. Dependent variables in Log. Robust standard errors in parentheses, clustered at the project level. Significance levels: 0.01: ***, 0.05: **, 0.10: *. Observations are weighted according to the inverse probability of treatment.*

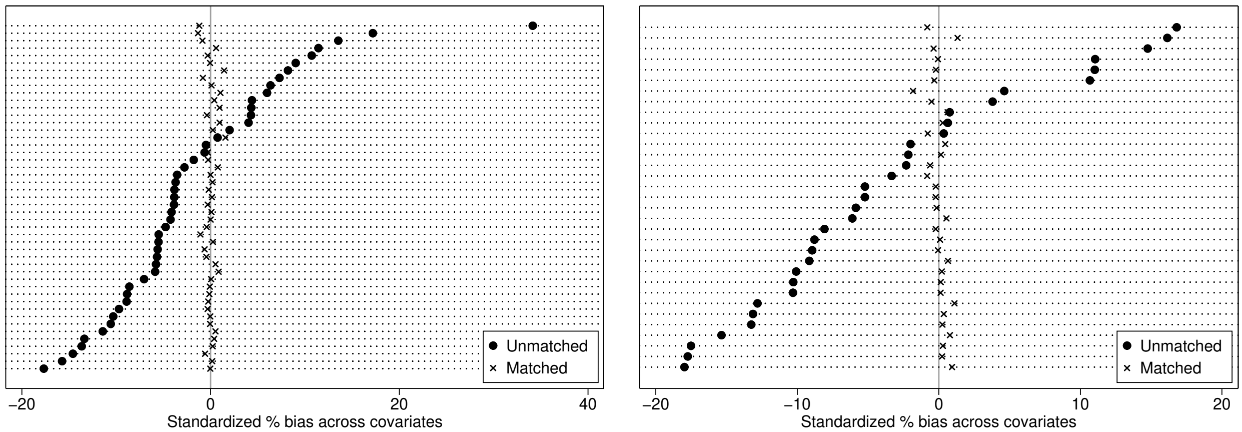
& Wahba (1999) suggest a balance test that builds upon this property. It consists in a comparison of weighted means between treated individuals and controls for each variable included in the logit regression.[17] All logit specifications have been initially retained because they pass this balance test. We present that test for our preferred propensity score estimation and weighting scheme in Figure 2.1. This figure shows that the mean differences are substantially reduced between the treated and controls due to reweighting, so that the remaining bias is close to zero for each covariate. This balance diagnosis test (standardized differences of the means tests for each covariate of the logit regression) is also performed within seven distinct strata of the propensity score (separately for the non-directed and directed programs because the preferred propensity score estimations are also performed separately), and our preferred model also passes that test.[18] Further, the treated and controls sets are also well balanced according to the propensity sores, when observations are weighted according to the IPTW method. Figures 2.2 and 2.3 show that matched samples are pretty similar in terms of propensity scores distribution for both directed and non-directed programs.

IV The impact of project funding

Figure 2.4 shows that ANR funds persons who have an increasing publication trend, which starts before the year of first funding (at $t=0$) and subsequently expands. Figure 2.4 also reports the properly weighted performances of their controls (dashed blue line). Note

[17] When possible, we use the procedure proposed by Becker et al. (2002).
[18] Appendix E is dedicated to balancing tests.

Figure 2.1 – Standardized bias (in %) associated with each explanatory covariate in the original unmatched sample and in the weighted sample for the directed (left graph) and non-directed (right graph) programs, using the estimated inverse probability of treatment weights.



Notes: Each dotted line represents an explanatory covariate included in the propensity score logit model (the X vector). The variables used in the propensity score model for non-directed programs are presented in Table 2.4, and those used for the non-directed programs are presented in Table 2.5.

those data are expunged of time trends and yearly shocks. We see that the performances of the controls are slightly lower than those of the treated individuals. This is because the pre-treatment difference-in-differences placebo test has given priority to the similarity in trends with the treated individuals. Other sets of controls which were more similar in outcome levels have been discarded because they do not satisfy all placebo parallel paths tests as well as the preferred ones.

And as expected, non-funded applicants have very similar trends to the funded agents until the year of funding (included). In fact, it turns out that the trends diverge only starting from the second year after funding. It is sometimes claimed that researchers have often nearly completed their project when applying. If these projects were also more likely to be funded, then a positive impact could be partially driven by this phenomenon. However, in that case, divergence should occur early after the funding date, something that is not observed here. This does not mean that anticipated projects are not more likely to be funded but that the conditions we imposed for the selection of controls seem to have sorted out such an effect.

Publication outcomes The main conditional difference-in-differences results are shown in Table 2.7. We find that receiving an ANR fund increases publications by 3.5% according to the preferred estimation. When the impact factor of the scientific journals in which articles are published is taken into consideration, receiving an ANR fund increases production by 8.3%. The impact of funding is strongest when citations are considered: a 15.2% increase is found. Impact is thus much stronger on indicators that capture the quality dimension of scientific output. The research project of the granted seem to attract more

Figure 2.2 – Density and box plot of the estimated propensity scores before and after weighting by the inverse probability of treatment weights for the directed programs.

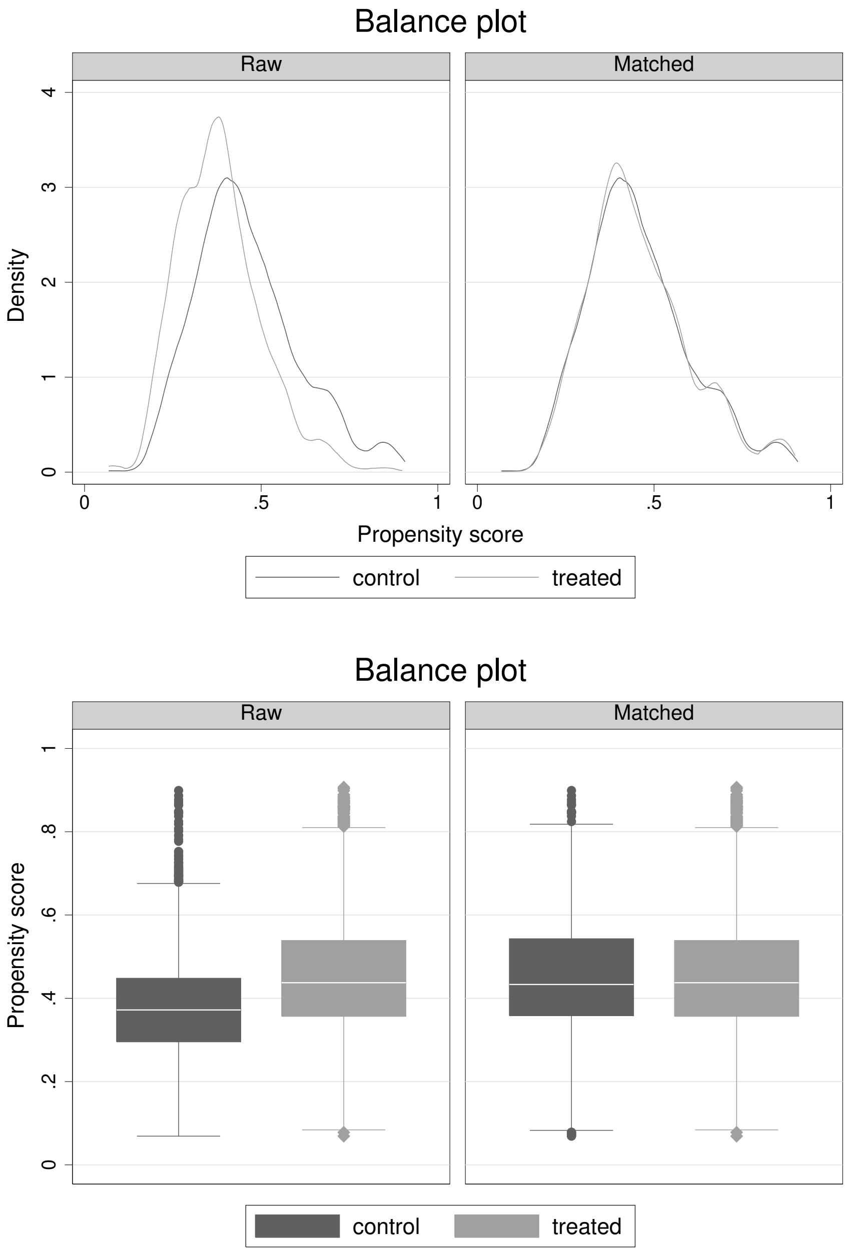


Figure 2.3 – Density and box plot of the estimated propensity scores before and after weighting by the inverse probability of treatment weights for the non-directed programs.

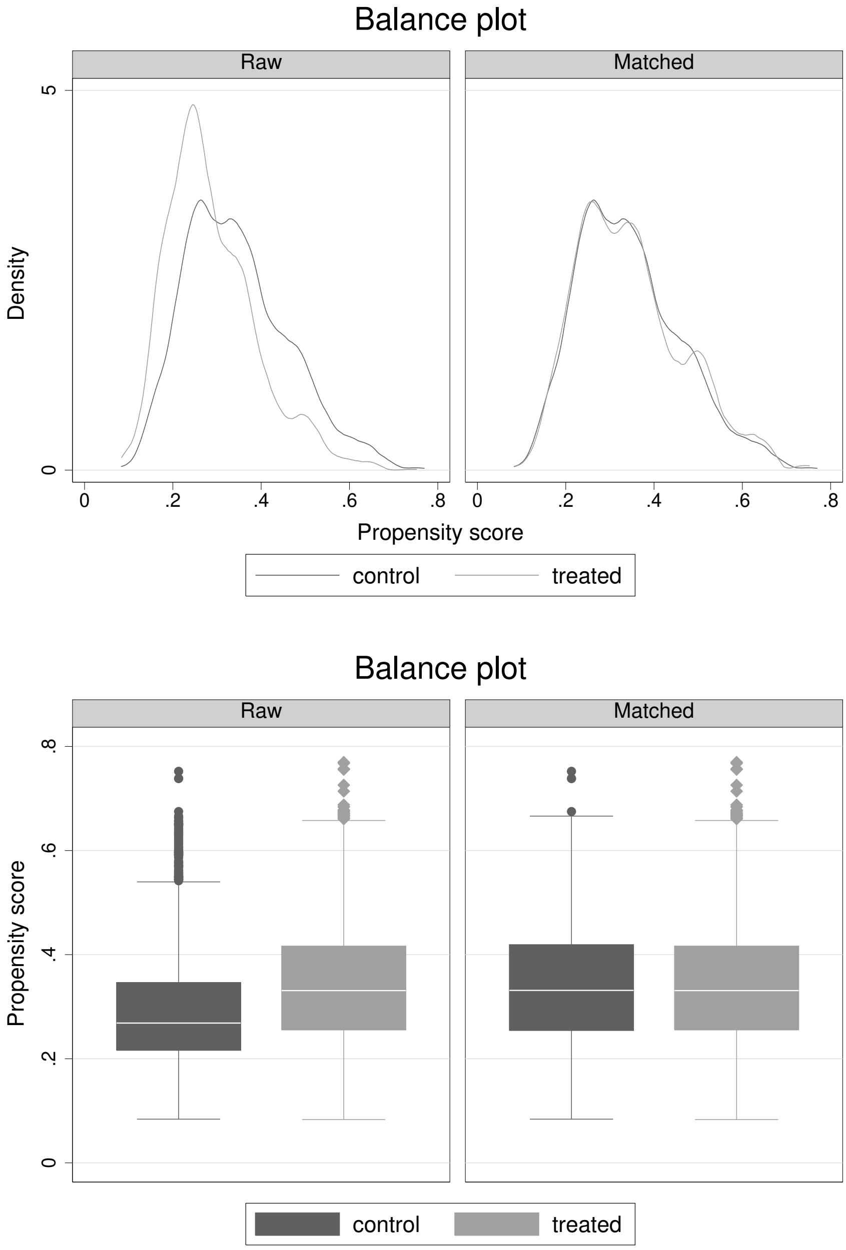
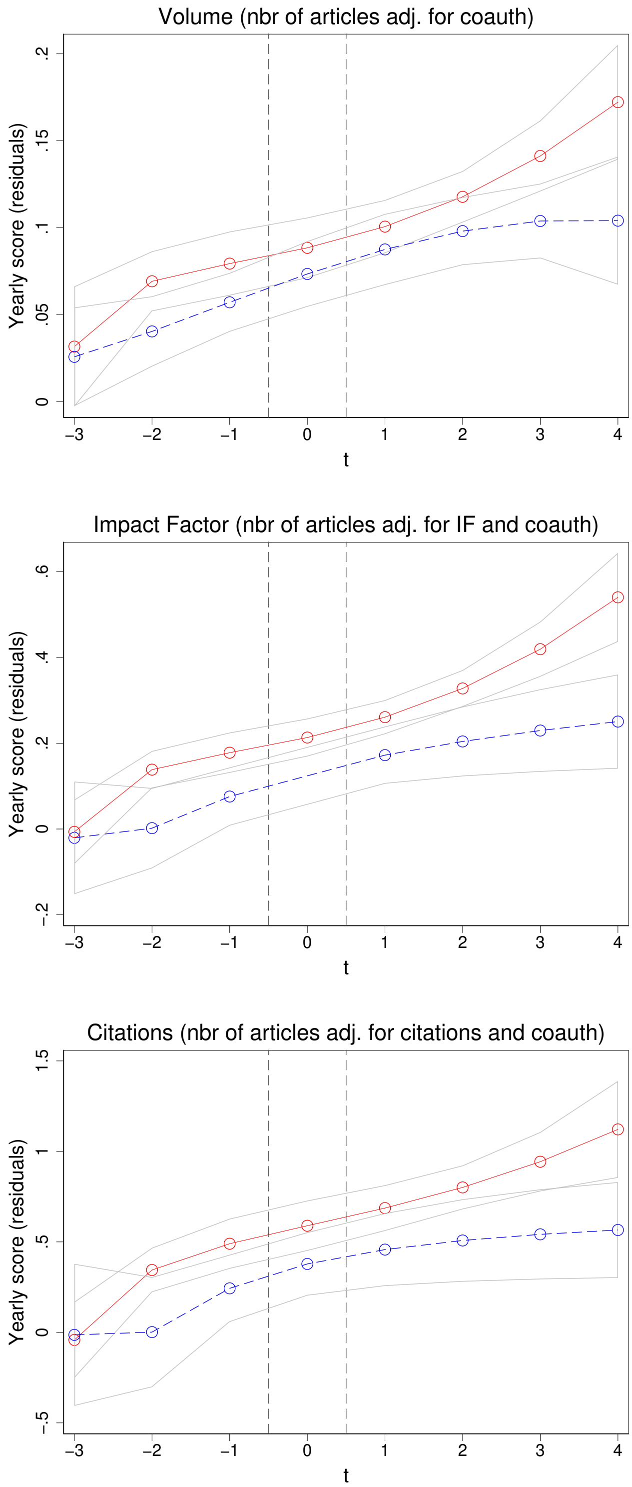


Figure 2.4 – Yearly scientific outcomes of the funded professors and researchers (red solid line) and of their controls (blue dashed line) with respect to the funding year ($t = 0$).



Notes: The red solid line stands for the granted and the blue dashed line stands for the unsuccessful applicants. Mean and 95% fractional polynomial confidence intervals are presented. The first year of funding occurs at $t = 0$. For each variable considered (Volume, Impact factor and Citations), we present the residuals obtained after regressing yearly scores on year dummies (absorbing potential year shocks and trends). Observations are weighted according to the inverse probability of treatment. The variables used in the propensity score models are reported in Tables 2.4 and 2.5.

attention from the scientific community which more likely cites their work.

As we have seen that age can influence scientific productivity, we are worried that these results may be slightly biased by age differences between treated individuals and controls. Moreover, the literature has long emphasized that age plays a significant role in scientific outcomes,[19] as an inverted-U shape of scientific productivity has been found in most fields of science. However, age differences between funded individuals and controls at the time of application are very limited in each funding program (see Table 2.2). Unreported regressions that are similar to the main ones but controlling for age and age squared exhibit no significant change in the results.

These results are larger than those obtained for NIH grants (7% impact on citations) by Jacob & Lefgren (2011), though the mean amount of the funds allocated in our sample is far less than the average NIH N01 grant. Jacob & Lefgren (2011) report a 1.7-million-US-dollar NIH N01 grant on average as compared to an average ANR grant of less than .14 million euros, and an average total cost of .56 million euros. This difference may be due to the specificity of the biomedical sciences in the US for which the availability of funds and the variety of funding sources may induce a displacement effect (as the authors themselves argue). Such an effect occurs if the funded individuals expend less energy in obtaining more funds than the unsuccessful applicants taken as controls. The plausibility of that explanation is reinforced by the fact that alternative sources of project funding than the ANR at the national level were relatively limited at the time of the study.[20] Note that our results are quite similar to those obtained in Gush et al. (2018), who use a different methodology, and data from a different country.

Collaboration patterns The literature has recently documented a long-run increase in the size of research teams proxied by the number of co-authors of the articles (Wuchty et al. 2007). We now document a hypothetical impact of project funding on team size. Coordinators may have incentives to delegate research tasks because they experience rising time constraints and because they have more financial resources to staff their teams. We find (see Table 2.7) a positive but limited impact of funding on the average number of authors per paper (2.2%). However, the impact of ANR funding on the total number of co-authors is significantly larger (9.8%). Thus, project-based funding increases the network of collaborators of the funded individuals more than it does the size of their research teams. This increase seems essentially due to the turnover of co-authors, as treated individuals have 6.7% more new collaborators than controls. This could be due to a higher capacity to hire PhD students or postdocs that eventually become co-authors on specific projects.

[19] To name a few: Lehman (1953), Zuckerman & Merton (1972), McDowell (1982), Levin & Stephan (1991).

[20] At the European level, the ERC was launched in 2007. It had however a limited budget in period 2007-2009: less than 1.7 billion euros for the whole of Europe. We matched the PIs of ERC grants in this period with our list of French professors and researchers, but found only a few scientists in the two lists.

It could also indicate that the funded individuals become more attractive as co-authors on the academic "collaboration market". To disentangle the two effects we would need to characterize further the collaborators of the treated individuals and controls, which is very difficult because of data limitations. We can however proxy the international span of their individual networks by counting the number of publications for which the authors gave at least one professional address outside France. Funding is found to increase the number of such articles by 4.2%, a result which is positive and significant albeit below the impact of funding on publication volume. This supports the idea that the two effects are at play: the funded individuals increase their networks by hiring, and also by collaborating more with independent colleagues. Moreover, this shows that ANR funding, which is mainly organized on a national basis, does not decrease the internationalization of collaborations but increases it, though to a limited extent.

Table 2.7 – Average treatment effect of receiving an ANR grant on publication outcomes and collaboration behaviors (the three years after treatment against the three years before).

Volume	Impact Factor	Citations
0.0350***	0.0825***	0.1525***
(4.46)	(7.53)	(9.30)
Av. Team Size	**Coauthors**	**Internat. Collab.**
0.0218***	0.0981***	0.0418***
(2.71)	(7.02)	(2.82)
New Coauthors[a]		
0.0668***		
(3.03)		

*Notes: Conditional difference-in-differences results. Coefficients and standard errors of the interaction term between the post-funding period dummy and the treatment dummy in fixed effect regressions. Observations are weighted according to the inverse probability of treatment. Dependent variables in Log. Robust standard errors in parentheses, clustered at the project level. Significance levels: 0.01: ***, 0.05: **, 0.10: *. The variables used in the propensity score models are reported in Tables 2.4 and 2.5.*

[a] *Conditional differences results only as this variable counts the new items in the post-treatment period as compared to the pre-treatment period for treated and control units.*

V The impact of directed vs. undirected programs

We now exploit variations in the program characteristics to uncover which funding design has larger impact on scientific outcomes. Over the considered period, the ANR ran two main types of programs: the directed and the undirected funding programs. The non directed programs are standard programs, open to any fields of science and managed by disciplinary based panels. The directed programs correspond to specific calls for project proposals in new fields of research for which the agency has diagnosed a specific need or opportunity for its financial support. The proposals are selected by transdisciplinary panels. Because these calls are targeted, only subsets of possible recipients can apply in practice. Therefore, non-directed programs are likely to be characterized by a higher degree of competition. Observed success rates are consistent with this statement: 37% in the directed programs vs. 25% in the non-directed ones. Therefore, self-selection is also likely to be more pronounced in the non-directed programs, and indeed (see Table 2.2) average applicants and funded via non-directed programs outperform on average applicants and funded to directed programs, when articles are weighted by their citations. Which of the two types of programs should be more efficient, in the sense that it has a larger impact on scientific outcomes? On the one hand, we expect that directed programs may make a big difference on targeted fields. If, as intended by the policy, it encourages the investigation of promising emerging research areas, it should lead to more path breaking research, leading to more cited papers published in well established journals. On the other hand, non-directed programs may have a larger impact because, thanks to a stronger competition and to their openness toward ideas heading in unspecified direction, they should be able to pick unexpectedly high quality projects.

Conditional triple difference model Our identification strategy builds upon the basic conditional difference-in-differences model by introducing a supplementary level of differentiation. As this basically differentiates double-differences, this estimation is called conditional triple-difference. We here shortly explain the model, before presenting results on the differentiated impacts according to the two types of programs launched, the directed (a non-neutral funding design) vs. the non-directed (a more standard and neutral funding design). For instance, the ATT differential of being treated by the non directed program as compared to being treated by a directed program is given by:

$$\begin{aligned}\hat{\delta}_{N-D} = {} & \frac{1}{|N_T^N|}\sum_{i\in N_T^N}\omega_i\left(Y_{i,1}-Y_{i,0}\right)-\frac{1}{|N_C^N|}\sum_{j\in N_C^N}\omega_j\left(Y_{j,1}-Y_{j,0}\right)\\ & -\left(\frac{1}{|N_T^D|}\sum_{i\in N_T^D}\omega_i\left(Y_{i,1}-Y_{i,0}\right)-\frac{1}{|N_C^D|}\sum_{j\in N_C^D}\omega_j\left(Y_{j,1}-Y_{j,0}\right)\right),\end{aligned} \tag{2.6}$$

where N stands for “non directed” or “neutral”, D stands for “directed”, N_T^p is the set of persons who received funding of type $p \in \{N, D\}$ and N_C^p is the set of controls for the funded individuals of type p. $Y_{i,t}$ is the outcome variable observed in period t, with $t = 1$ in the period after the treatment assignment, and $t = 0$ in the period before treatment. The weights ω_j are defined as in Equation 2.5. The first part of the right side of the equation refers to the difference between the treated and control groups of non directed programs, whereas the second part is the same difference for directed programs. The differential ATT of non directed programs over directed programs is simply equal to the difference between those two terms. It is estimated using a similar regression as Equation 2.4, but now considering the coefficient of a term to be added, formed of a triple interaction between a post-funding dummy, a treatment dummy and a non-directed program dummy.

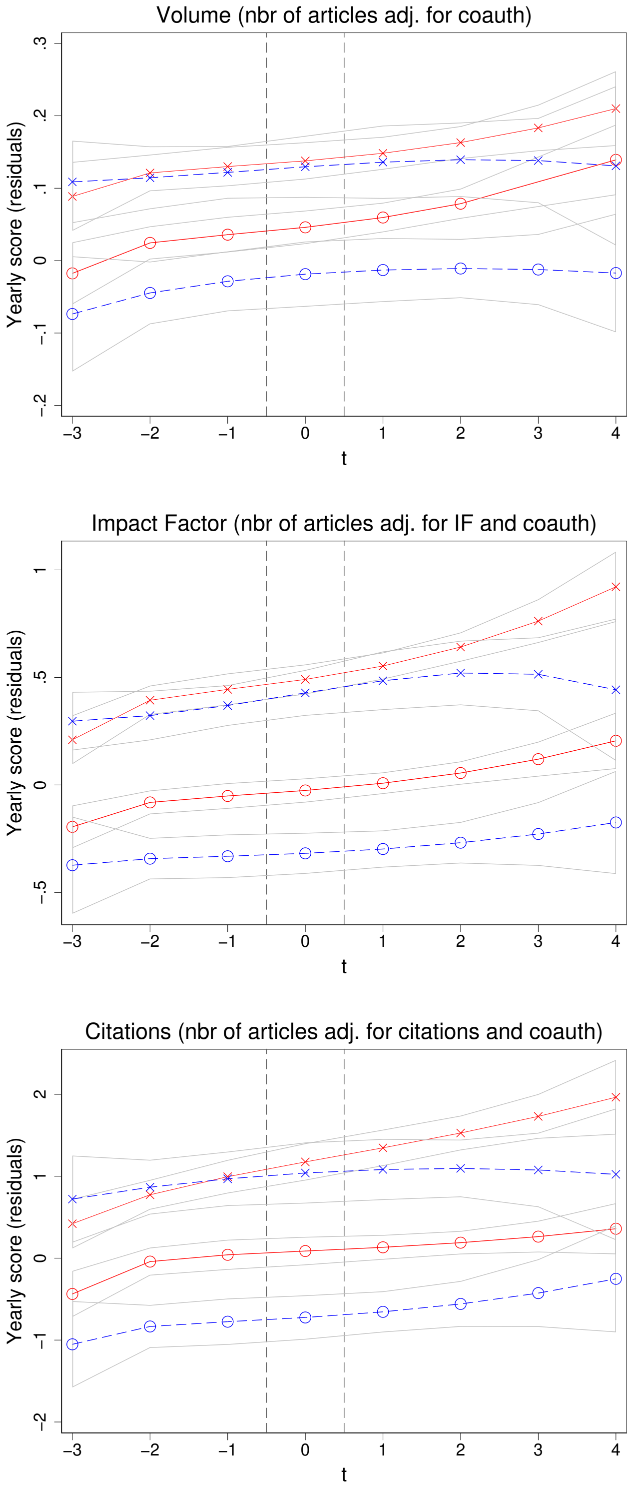
The impact of directed vs. non-directed programs Figure 2.5 shows the publication records of the granted and the (properly weighted) applicants to the two type of programs at different years before and after the application year. Red solid (blue dashed) lines stand for the granted (controls). Circles (crosses) denote directed (non-directed) programs. Controls in each program have publication trends that are very similar to the granted before granting date. For non directed programs, treated and controls match nearly perfectly in levels as well before treatment. Applicants and granted to the non-directed programs have larger publication records than their counterparts in the directed programs which is consistent with the idea of a higher level of competition in this program. As a first sign of a presumed superior efficiency of non-directed programs, we observe an increasing spread between granted and controls posterior to the application year, for this program only when publications are weighted by citations or by the journal Impact Factor.

The precise impact analysis is reported in Table 2.8. Directed and non-directed programs barely differ in their impact on the volume of scientific production: a 2.8% difference in favor of non-directed programs, only significant at the 10% level. However, non-directed programs turn out to be significantly more efficient when the impact factor of the journals or the number of direct citations are taken into account. Directed programs have a treatment effect on the treated (baseline ATT in the table) of 3.1% when articles are weighted by the journal Impact Factor, while switching to a non-directed program increases that outcome by 11.1%. The difference between directed and non-directed programs is even sharper when articles are weighted by citations: the baseline treatment effect of directed programs is 5.9%, while switching to a non-directed program raises output by a 20.3%. These differences between program types are even larger than the overall impact of ANR funding.

These results strongly support the idea that non-directed programs are very efficient, while directed programs have a limited impact on scientific outcomes.

Novelty Peer review procedures have been repeatedly criticized as being negatively biased toward really groundbreaking and innovative projects (Braben 2004, Chubin et al. 1990, Wessely 1998). Boudreau et al. (2016) show that highly novel projects are associated with lower ratings in a field experiment. Azoulay et al. (2011) show that scientists supported on a program specifically funding researchers (vs. projects) explore more novel research lines. The authors interpret their finding arguing researchers granted on projects

Figure 2.5 – Yearly scientific outcomes of the funded professors and researchers (red solid line) and of their controls (blue dashed line) who applied to the two funding schemes: directed (o marks) and non-directed ($\times$ marks) with respect to the funding year ($t = 0$).



Notes: The red solid line stands for the granted and the blue dashed line stands for the unsuccessful applicants. The circle points correspond to directed programs while the crosses stand for the non-directed programs. Mean and 95% fractional polynomial confidence intervals are presented. The first year of funding occurs at $t = 0$. For each variable considered (Volume, Impact factor and Citations), we present the residuals obtained after regressing yearly scores on year dummies (absorbing potential year shocks and trends). Observations are weighted according to the inverse probability of treatment. The variables used in the propensity score models are reported in Tables 2.4 and 2.5.

Table 2.8 – Differentiated effects of receiving an ANR grant on outcomes according to non-directed versus directed funding schemes (the three years after treatment against the three years before). Baseline average treatment effect for directed programs are in italics.

	Volume	Impact Factor	Citations
Non-Directed vs. Directed programs	0.0277* (1.77)	0.1111*** (5.10)	0.2028*** (6.26)
Baseline ATT of Directed programs	*0.022*** *(1.99)*	*0.0314*** *(2.04)*	*0.059*** *(2.54)*
	Av. Team Size	**Coauthors**	**Internat. Collab.**
Non-Directed vs. Directed programs	-0.0011 (-0.07)	0.0201 (0.72)	0.0286 (0.97)
Baseline ATT of Directed programs	*0.0223*** *(2.36)*	*0.0885**** *(4.48)*	*0.0283* *(1.32)*
	New Coauthors[a]		
Non-Directed vs. Directed programs	-0.049 (-1.12)		
Baseline ATT of Directed programs	*0.0901**** *(2.91)*		

*Notes: "Non-Directed vs. Directed programs" lines report conditional difference-in-difference-in-differences results. Coefficients and standard errors of the triple interaction term between the post-funding period dummy, the treatment dummy and the non-directed-program dummy, in fixed effect regressions. "Baseline ATT of Directed programs" lines report the estimation of Conditional difference-in-differences for the directed programs only. Coefficients and standard errors of the interaction term between the post-funding period dummy and the treatment dummy in fixed effect regressions. All observations are weighted according to the inverse probability of treatment. Dependent variables are in log. Robust t-stats in parentheses, clustered at the project level. Significance levels: 0.01: ***, 0.05: **, 0.10: *. The variables used in the propensity score models are reported in Tables 2.4 and 2.5.*
[a] *Conditional differences results only as this variable counts the new items in the post-treatment period as compared to the pre-treatment period for treated and control units.*

are bound to their project proposals whereas others can more easily redesign their goals. Agencies could also face more difficulties in inducing (often disciplinary based) committees to support risky research projects rather than to fund researchers willing to take such risks. As we have access to two different types of funding programs, we can specifically look how both programs deal with novelty. Directed programs focus primarily on new and promising areas of science. If successful in their explicit goals, they should attract and fund more often professors and researchers who investigate new research problems. However, the non-directed programs, which are open to any field of science, may as well attract and select unanticipated pathbreaking research proposals.

To address this issue, we need to proxy the novelty of the research articles of funded and non successful applicants to the directed and non-directed programs, before and after funding date. As we need to look at this dimension in the longer run, we perform a supplementary extraction of WoS data up to year 2015 that is up to six years after the last funding year of funding (2009). Article novelty is calculated using the frequencies of pairwise combinations of Author Keywords as introduced in Carayol et al. (2018). This measurement of novelty is intended to identify the originality of the research directions, the very problem addressed by research articles. Carayol et al. (2018) show, on more than ten million research articles published by journals indexed in the Web of Science (WoS), that Pairwise Author Keywords Novelty is a very good predictor of citations and highly cited articles, even in the relatively short run.[21] We compute yearly average and maximum pairwise author keywords novelty to appreciate to what extent their research is novel over time, before and after the application year.

In Figure 2.6, as in previous figures, we use residuals obtained after yearly scores are first regressed on year dummies. We find that directed programs indeed attract and fund professors and researchers whose research is more novel on average than non-directed programs. Differences between programs are more pronounced when we look at maximum novelty rather than at average novelty. It is interesting to observe that granted professors and researchers on non-directed programs perform less novel research than unsuccessful applicants. This difference however shrinks when considering maximum novelty. There is no post-treatment tendency of the granted from directed programs to specifically undertake more novel research. If significant, the impact would rather be negative but Table 2.9 confirms there is no overall significant impact of funding on novelty, and no significant differentiation between programs in this respect.

These results lead to the conclusion that directed programs are more successful in attracting and funding researchers and professors who produce more novel science. However, both directed and undirected programs are ineffective in incentivizing the funded toward addressing more novel research lines than they did before treatment, even in the longer run.

VI Designing funding programs: more results

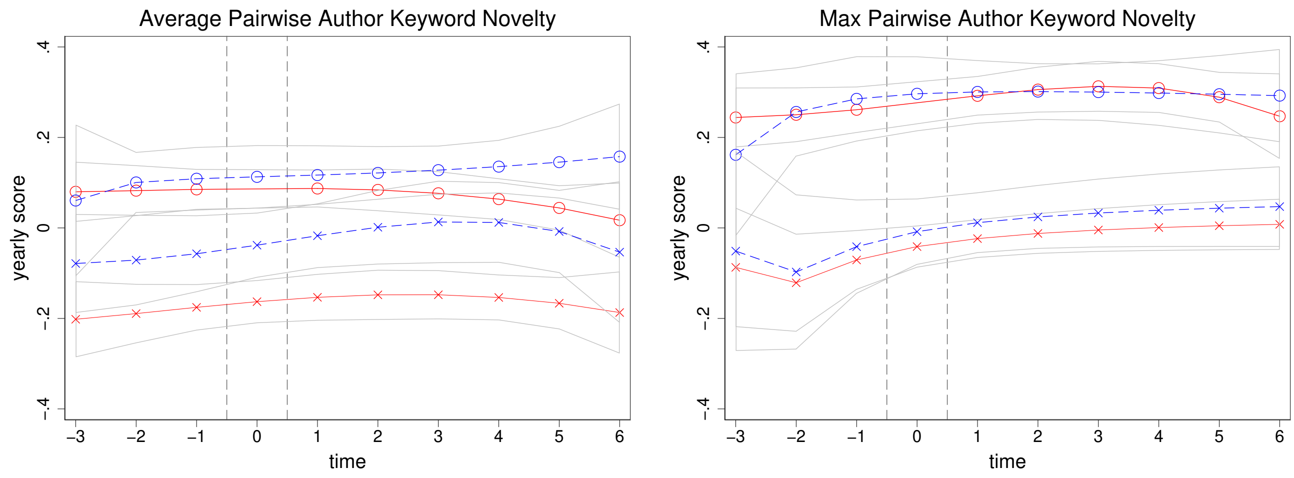
We want to shed light on the conditions under which project funding turns out to be more efficient, and to what extent precisely.[22]

Impacts along the career path Estimating the impact of fund allocation at different career stages is an important policy issue. We thus ran estimations similar to the preceding ones, allowing us to differentiate the impact on younger scientific coordinators (equal to or less than 43 years old, the median age) from that on older ones. The impact of choosing a

[21] Carayol et al. (2018) show that relying on the frequencies of pairs of author keywords is key to this result, as using either keyword frequencies, predefined keywords or journal co-citations does not lead to the same results.

[22] The main results are presented in this section whereas some of their associated tables are to be found in Appendix F.

Figure 2.6 – Yearly average (left graph) and maximum (right graph) Pairwise Author Keyword Novelty of professors and researchers' research articles. Red solid lines (blue dashed line) stand for the funded (unsuccessful applicants). Applicants to directed programs (non directed programs) have *o* marks (× marks).



Notes: The red solid line stands for the granted and the blue dashed line stands for the unsuccessful applicants. The circle points correspond to non-directed programs while the crosses stand for the directed programs. Mean and 95% fractional polynomial confidence intervals are presented. The year of first funding occurs at $t = 0$. The included data points go up to year 2015 included, that is up to six year after the last funding year of funding (2009). For each variable considered (mean and maximum article novelty in the considered year), we present the residuals obtained after regressing yearly scores on year dummies (absorbing potential year shocks and trends). Observations are weighted according to the inverse probability of treatment. The variables used in the propensity score models are reported in Tables 2.4 and 2.5.

younger coordinator is then estimated by interacting post-funding dummy with treatment dummy in a fixed effect regression using two time periods' panel data, where observations are weighted according to the chosen method. Results are reported in Table 2.10. We find non significant differences in the volume of publications and when articles are weighted by the journal's average impact factor. However, an important and significant difference is found in terms of citations: the impact on younger coordinators is 9.5% higher than that on older ones. This implies that the impact in terms of citations for younger coordinators is more than twice that observed among older scientists. This result is pretty strong and has significant policy implications. Further, no significant differentiated effect on collaborations is observed. Funding only increases the team size of the older scientists slightly more than that of the younger scientists (2.8%, significant at the 10% level only).

We now differentiate the impact according to the publication profiles of the treated individuals at the time of funding. Our goal is to investigate whether some publication profiles are more likely to be positively impacted by the funding policy than others. Treated individuals and controls are ranked within each discipline according to the number of citations received by their articles published in the preceding three years,[23] and are categorized in either one of the four largest deciles or in the remaining six deciles. In the triple difference approach, the performances of the top 10% are taken into reference. It

[23]We have used alternative performance variables to rank them, such as the number of articles, or even when such articles are weighted by the journal impact factor. Results are qualitatively similar.

Table 2.9 – Differentiated effects of receiving an ANR grant on the average and maximum Pairwise Author Keywords Novelty according to non-directed versus directed funding schemes (the three years after treatment against the three years before). Average treatment effect for the baseline (all projects and directed programs) in italics.

	Average Pairwise Author Keywords Novelty	Maximum Pairwise Author Keywords Novelty
Non-Directed vs. Directed programs	-.0071 (-0.56)	-.0097 (-0.76)
Baseline ATT	*-.0038* *(-0.69)*	*-.005* *(-0.83)*
Baseline ATT of Directed programs	.0027 (0.29)	.0019 (0.21)

*Notes: "Non-Directed vs. Directed programs" lines report conditional difference-in-difference-in-differences results. Coefficients and standard errors of the triple interaction term between the post-funding period dummy, the treatment dummy and the non-directed-program dummy, in fixed effect regressions. "Baseline ATT of Directed programs" lines report the estimation of Conditional difference-in-differences for the directed programs only. Coefficients and standard errors of the interaction term between the post-funding period dummy and the treatment dummy in fixed effect regressions. All observations are weighted according to the inverse probability of treatment. Dependent variables are in log. Robust t-stats in parentheses, clustered at the project level. Significance levels: 0.01: ***, 0.05: **, 0.10: *. The variables used in the propensity score models are reported in Tables 2.4 and 2.5.*

is found that the treated individuals who are in the top 10% are never those on which the impact is the largest. Largest impacts are found in terms of publication volume when the treated individuals are in the second to the fourth deciles only, which are significantly larger than those of the first decile (from 8.2% to 11.8% larger). Similar statements can be made in terms of impact factor and citations, though coefficients are less significant. This can be explained by the fact that the top professors and researchers may have access to other sources of funds. Though the committees should select applicants who have strong publication records, the impact is not likely to be the largest when the funds are targeted to those who can obtain funds elsewhere, at the European level, for instance. Note that this statement is in terms of elasticities, not in absolute terms (number of citations for instance). A lower impact in terms of elasticity on top-10% performers may well correspond to a larger impact in absolute outcomes. On the other side of the distribution, when the treated individuals are not in the four largest deciles, the impact is likely to be significantly lower, not on the volume of publications, but both when the impact factor of the journal is considered and for citations. When, for instance, in the six lower deciles, the treated individuals have an average impact in terms of citations reduced by 9%, that is no longer significantly different from zero.

PI or not PI Project variables are also available. In particular we have information on the role each person plays in the project: is she/he scientific coordinator of the whole project (the PI of the project), or only scientific coordinator of one institutional partner in a multi-partner project. As the design of the ANR grant system provides each partner's

Table 2.10 – Differentiated effects of receiving an ANR grant on outcomes according to age dummy: below the median age (43) versus over the median age (the three years after treatment against the three years before).

	Volume	Impact Factor	Citations
Young (below median age) vs. Older	0.0221 (1.41)	0.0266 (1.29)	0.0952*** (3.09)
Baseline ATT on the Older	0.0227** (2.00)	0.0679*** (4.49)	0.1008*** (4.45)
	Av. Team Size	**Coauthors**	**Internat. Collab.**
Young (below median age) vs. Older	-0.0279* (-1.80)	-0.0034 (-0.13)	0.0280 (0.95)
Baseline ATT on the Older	0.0367*** (3,31)	0,0993*** (5,02)	0,0266 (1,27)
	New Coauthors[a]		
Young (below median age) vs. Older	-0.0568 (-1.37)		
Baseline ATT on the Older	0.0994*** (3.17)		

*Notes: "Young (below median age) vs. Older" lines report conditional difference-in-difference-in-differences results. Coefficients and standard errors of the triple interaction term between the post-funding period dummy, the treatment dummy and the below median age dummy, in fixed effect regressions. "Baseline ATT on the older" lines report the estimation of Conditional difference-in-differences for the directed programs only. Coefficients and standard errors of the interaction term between the post-funding period dummy and the treatment dummy in fixed effect regressions. All observations are weighted according to the inverse probability of treatment. Dependent variables are in log. Robust t-stats in parentheses, clustered at the project level. Significance levels: 0.01: ***, 0.05: **, 0.10: *. The variables used in the propensity score models are reported in Tables 2.4 and 2.5.*
[a] Conditional differences results only as this variable counts the new items in the post-treatment period as compared to the pre-treatment period for treated and control units.

scientific coordinator with a significant level of autonomy (in particular financial), we have chosen the partner level of analysis rather than the project level. However, the project PI role is specific, often not a desirable one to play and one that keeps busy with administration and coordination tasks. We thus keep track of the status of each partner's scientific coordinator in the project with a dummy labeled PI, which will allow us to check whether PIs are compensated for their efforts by increased scientific productivity and/or collaborations.

In a project, do partners free-ride on the PI who bears most of the between-partners coordination costs? Or, conversely, does the project PI free-ride on the partners' scientific coordinators, using their labor force to increase his or her scientific production? We find no significant difference according to the status of the treated individuals in the project, who can be either PI or partner scientific coordinator. Thus it seems that the benefits and costs of coordinating multi-partner projects counterbalance each other. Gains of assuming

Table 2.11 – Differentiated effects of receiving an ANR grant on publication outcomes according to the position in the citation distribution at the time of funding (the three years after treatment against the three years before).

	Volume	Impact Factor	Citations
Baseline Top-10% publication performance	-0.0474**	-0.0340	0.1037**
	(-2.27)	(1.15)	(2.29)
Top-10-to-20% publication performance (vs. top-10%)	0.0823***	0.0608	0.0989*
	(3.06)	(1.62)	(1.80)
Top-20-to-30% publication performance (vs. top-10%)	0.106***	0.0675*	0.0716
	(4.05)	(1.87)	(1.33)
Top-30-to-40% publication performance (vs. top-10%)	0.118***	0.0880**	0.0921*
	(4.42)	(2.39)	(1.68)
Bottom-60% publication performance (vs. top-10%)	0.0632**	-0.0188	-0.0898*
	(2.38)	(-0.55)	(-1.70)

*Notes: "Baseline Top-10% publication performance" report the estimation of Conditional difference-in-differences for the top 10% publishing professors and researchers only. Coefficients and standard errors of the interaction term between the post-funding period dummy and the treatment dummy in fixed effect regressions. The other lines report the conditional difference-in-difference-in-differences results. Coefficients and standard errors of the triple interaction term between the post-funding period dummy, the treatment dummy and the percentile-class-of-the-citations-volume-prior-to-application dummy (mentioned at the right of each line, the top-10% being in reference), in fixed effect regressions. All observations are weighted according to the inverse probability of treatment. Dependent variables are in log. Robust t-stats in parentheses, clustered at the project level. Significance levels: 0.01: ***, 0.05: **, 0.10: *. The variables used in the propensity score models are reported in Tables 2.4 and 2.5.*
[a] *Conditional differences results only as this variable counts the new items in the post-treatment period as compared to the pre-treatment period for treated and control units.*

the PI role are also not observed in collaborations. Unreported estimates show that the PI role has no effect on team size, number of co-authors and number of new co-authors. These results highlight how burdensome the PI role is. At the time of the proposal, assembling together all partners' contributions. At the time of the project, coordinating the work of all partners. The specific rules of the ANR, which give broad autonomy to the institutional partners and thus less power to the PI of the projects, probably does not help reduce such coordination costs. Another explanation is that it is still complex (though not impossible) in France to use project funds to reduce coordination costs or at least buy back teaching time, for instance.

Year effect No significant difference is found according to the year of funding. This result may seem surprising, bearing in mind that the agency was created in 2005. We guess that the agency has significantly increased its capabilities over the time period considered. We also know that the level of competition has been fairly different across years. For instance, the rate of success of the first year was much higher than that of the second year (48% as compared to 26%). In a sense, the fact that we find no significant difference between years is reassuring vis-à-vis our estimation methodology - tending to show that appropriate controls have been found for each year.

Scientific field effect When interaction with the scientific discipline is considered, we find that the impact of receiving funds is never significantly larger than in the life sciences, which is the reference. The only exception applies to Information and Communication Sciences and Technologies, where the impact is greater by 8.8% on citations and by 6.2% on the number of articles. Note that significance levels are however low (in particular for citations) and should thus be treated with caution.

VII Conclusions and discussion

In this chapter we have taken advantage of the recent French experience in which a new institution for project-based funding was created in 2005. This institution operates on a large scale, having distributed funds to research projects whose accumulated total costs approach ten billion euros over the five years covered by the study. Moreover, a certain level of variation in programs' rules and recipients' characteristics allows us to investigate the relative efficiency of variants of project funding. The results are not specific to one field of science, as all disciplines of hard and natural sciences are concerned (as well as some social sciences).

We identify the impact of receiving a research grant essentially by comparing the research production trajectories of the scientific coordinators of the funded projects with those of control groups. The controls are selected and weighted thanks to propensity scores that model the treatment allocation on observables. Because the data on the whole reference population (not only on applicants) as well as several useful variables potentially explaining selection are available, we are able to define eight sets of controls based on different specifications of the propensity score model, which all pass the balance test. We compare how these various sets of controls pass parallel paths tests. The "best" set of controls according to those tests picks controls among applicants exclusively, models treatment by program types, and includes past publication performances at the time of treatment as well as recent trends. This suggests future studies should have similar information to obtain satisfactory control sets.

Concerning the global efficiency of project-based funding, our study concludes that a grant increases the number of publications weighted by citations by about 15%. That result is larger than what was previously observed in Jacob & Lefgren (2011). However, as our study is not limited to a specific scientific field and as few alternative opportunities for project-based funding were available at the time of the study in France, our results are less prone to be affected by a displacement effect (negative bias). This suggests that our quantification of project-based funding is the closest to the real effect.

Further, we also find that funding has a positive effect on the size of collaborators' network and on the turnover of collaborators. Although the agency under investigation operates on a national basis, it does increase international collaborations. Funding thus has a significant and positive impact on the scope of collaboration networks. One concern remains, however, since project funding does not affect the novelty of the research problems that are tackled by the funded individuals. This is a serious issue often raised by funding agencies themselves which would need further investigation.

Some of the most striking practical results of our study concern the differentiated impacts with respect to the types of program. We find that when programs have no specific direction, so that they are open to wider competition, they have a much larger impact. Directed programs have a significant but rather small impact, while the surplus of impact of non-directed programs is quite large, even larger than the average impact of funding. This nominal advantage of non-directed programs is not counterbalanced by any sign of increased novelty of the research performed by recipients of directed grants. However, the directed programs prove successful in attracting and funding professors and researchers who develop (essentially before the funding date) more novel research than non-directed programs. Last but not least, the funds allocated to younger applicants have much larger impacts than those allocated to older applicants. This strongly supports the idea that project-based funding should keep a large door open to younger applicants. If confirmed by other studies, these results may provide some guidelines for improving project funding in science.

Appendix A. Descriptive statistics

Out of our sample of 31,081 professors and researchers, 10,722 are identified as applicants to at least one ANR application during the period 2005-2009. Among these applicants, 5,786 were awarded a grant at least once (see Table 2.12).

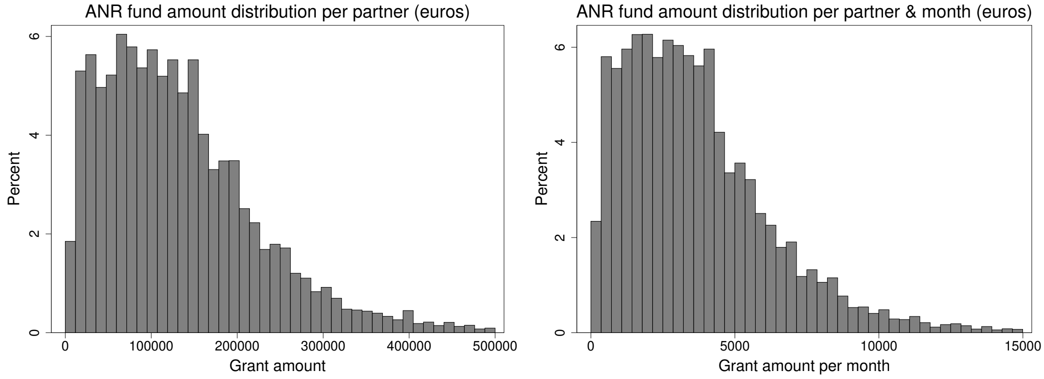
Table 2.12 – The final sample of 31,081 researchers, the applicants and the granted

	number	share
Non applicants	20,359	65.50
Applicants	10,722	34.50
Granted	5,786	18.62
Total	31,081	100.00

Descriptive statistics on the whole set of applications

The left graph of Figure 2.7 shows the distribution of ANR grant amounts assigned to all the funded projects in the period 2005-2009. The distribution is a log-normal shape, with a plateau around the median (equal to 136,000; the mean is 180,000) and a long right tail. The shape of the monthly subvention distribution (right graph of Figure 2.7) is similar to the previous one, with a mean of 4,800 and a median of 3,700.

Figure 2.7 – Histogram of the amount of ANR funding



Notes: The right tail of the distribution is cut to preserve confidentiality.

The yearly number of applications, number of fundings, total amount of the grants attributed and total cost of the subsidized projects are presented in Table 2.13. The selection rate is 30% on average, which leads to an accumulated grant amount of 2.4 billion euros over the period 2005-2009. The selection rate was significantly higher in 2005, due to a lower number of applications in the first year of agency operation. The estimated total cost of the funded projects is almost four times higher than the amount of the bestowed ANR grant.

Table 2.13 – Applications from the public sector and funded partner×project by years (in # and amounts in million euros)

-	# Applications	# Granted	Grants amounts	Total Cost
2005	5,616	3,553	417 me	1,510 me
2006	12,881	4,188	511 me	2,160 me
2007	11,655	3,496	483 me	1,910 me
2008	9,769	3,315	522 me	2,040 me
2009	14,931	2,890	470 me	1,890 me
Total	54,852	17,442	2,403 me	9,510 me

Descriptive statistics on the final sample

The yearly number of applications and amount of funding granted are presented in Table 2.14. The matching between the administrative list of French researchers and the ANR applications data set allowed for the proper identification of 46.2% of the applications and 45.5% of the granted ones. The related ANR budget represents 46.9% of its total outlay over the period.

The age distribution of the researchers and professors is presented for the three samples in Figure 2.8. We can see therein that the distributions are quite similar, although the 35 to 50 years old are, in proportional terms, slightly more numerous in the samples of applicants and grantees compared with the overall population.

In Table 2.15, we can see that full time researchers (denoted by Assist. Resear. and Research Dir.) and full professors (Full Prof.) are proportionally over-represented among the applicants. Senior researchers (Research Dir.) tend to be even more represented when considering the allocated grants. This is mainly due to the participation and success of CNRS researchers, who represent the vast majority of researchers in our database (see Table 2.16). Although assistant professors and full professors (72,2% of the population, denoted by the "UNIV" acronym) prevail in our sample, this group has the lowest share of funded individuals: 1.4 granted out of 10 individuals, whereas the other groups have at least 3 granted out of 10 individuals (with the exception of IRD).

Table 2.14 – Applications from the public sector and funded partner×project by year for our final sample (in # and amounts in million euros)

-	# Applications	# Granted	Grants amounts	Total cost
2005	5,422	2,605	185 me	476 me
2006	8,072	2,125	237 me	1,100 me
2007	4,994	1,473	243 me	984 me
2008	3,559	1,000	231 me	985 me
2009	3,317	742	232 me	856 me
Total	25,364	7,945	1,128 me	4,400 me
Prop	46.2%	45.5%	46.9%	46.3%

Notes: The total number of applications is higher than the number of applicants (10,722) because they applied 2.37 times on average over the period. Some of the applicants also received multiple fundings (this accounts for 5,786 applicants granted).

Figure 2.8 – Age histograms for the total population, the applicants and those funded

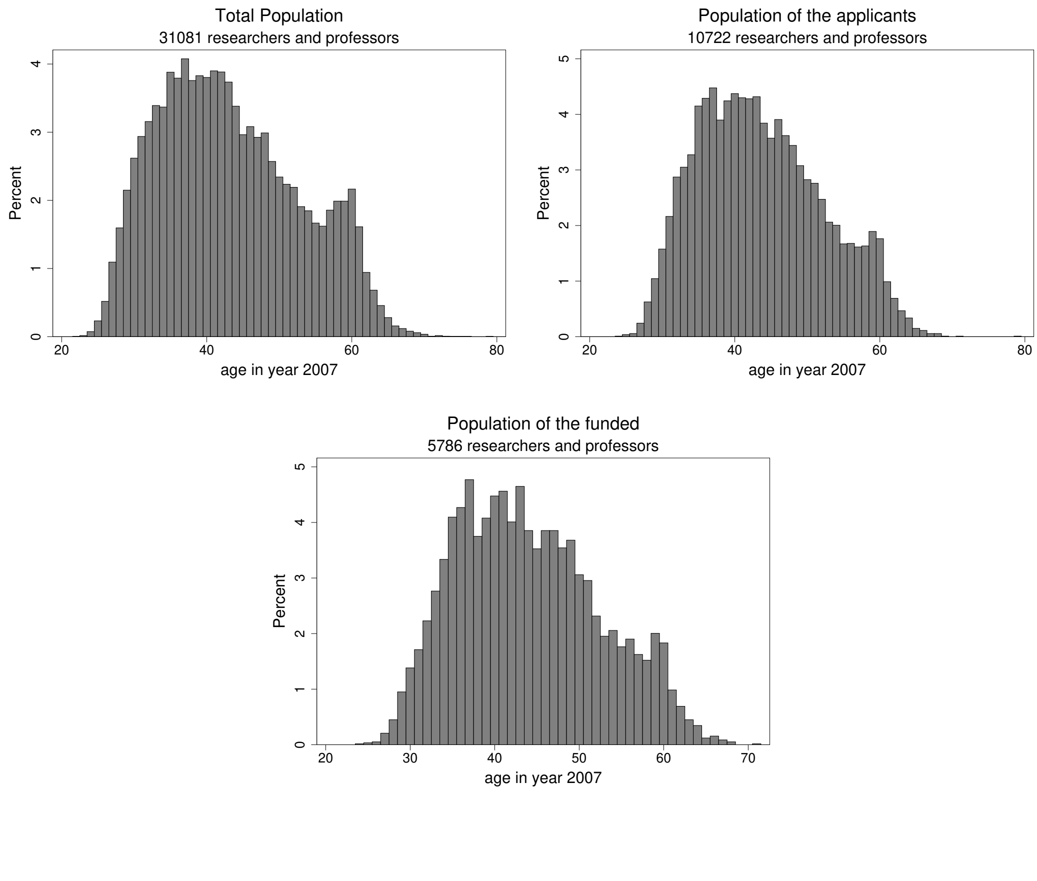


Table 2.15 – Researchers' and professors' status in the three samples

-	Full sample		Applicants		Grantees	
	#	%	#	%	#	%
Assist Resear.	5,290	17.02	2,335	21.78	1,260	21.78
Research Dir.	3,340	10.75	2,153	20.08	1,462	25.27
Assist Prof.	13,887	44.68	2,679	24.99	1,115	19.27
Full Prof.	8,564	27.55	3,555	33.16	1,949	33.68
Total	31,081	100.00	10,722	100.00	5,786	100.00

Notes: "Assist Resear." is assistant researcher, whereas "Research Dir." stands for research director. These two statuses represent positions that are dedicated full-time to research activity. In France, all these statuses confer a civil servant position and therefore imply tenure.

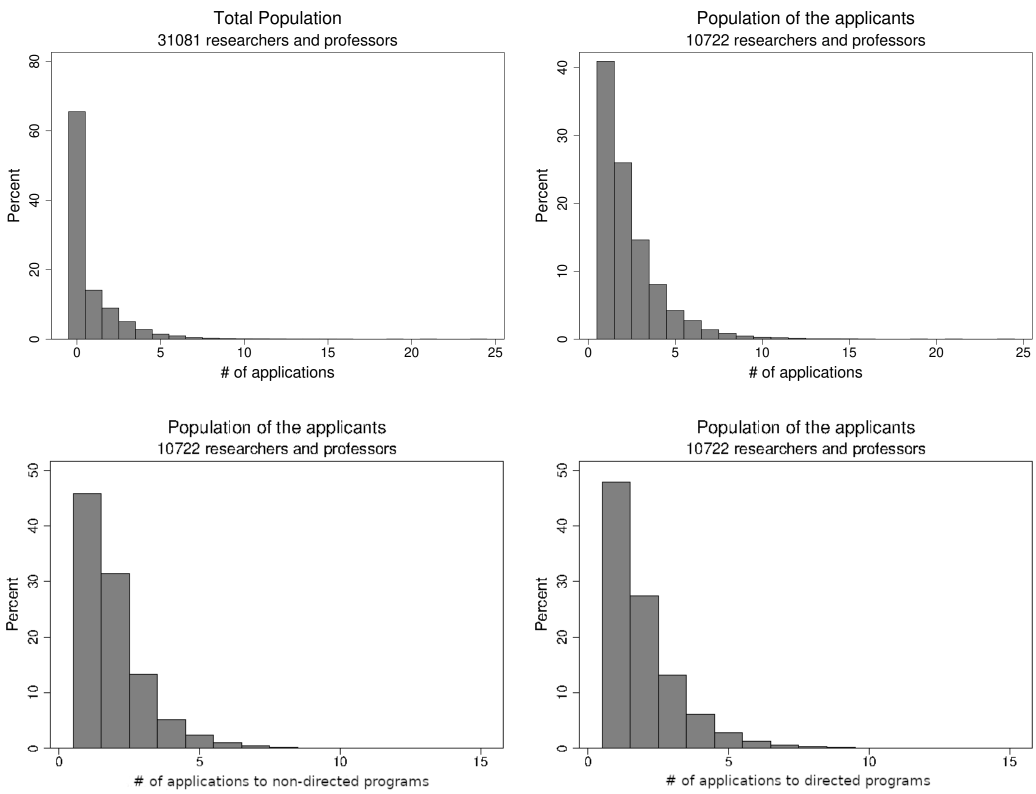
Table 2.16 – Researchers and professors' employing institutions in the three samples

-	Full sample		Applicants		Granted	
	#	%	#	%	#	%
CNRS	6,580	21.17	3,473	32.39	2,114	36.54
INRA	380	1.22	185	1.73	113	1.95
INRIA	146	0.47	82	0.76	58	1.00
INSERM	1,290	4.15	668	6.23	396	6.84
IRD	235	0.76	81	0.76	41	0.71
UNIV	22,450	72.23	6,233	58.13	3,064	52.96
Total	31,081	100.00	10,722	100.00	5,786	100.00

Notes: "UNIV" stands for universities. The CNRS is a public institution, which supports research in any scientific field. The remaining public organizations are specialized ones: the INRA is the national agronomic research institute, the INRIA is the national research institute of computer science and automation, and the INSERM is the national institute for health and medical research. The IRD is the national institute for development.

In Figure 2.9, we see that the distribution of the number of applications is skewed to the right, with most professors and researchers not applying. Among the ones who apply, most apply only once, while some are applying many times. The applications in our final sample are equally divided between directed and non-directed programs on the whole period (see Table 2.17). In the first years after the creation of the ANR, the number of applications to directed programs is higher. The importance of the two types of programs gradually balances before reversing in 2008. In 2009, non-directed applications represent nearly two thirds of all applications. When we look more in detail at the applications to the seven specific directed programs that were launched,[24] we observe that the number of applications is highest for the Biology and Health program. We also note significant variations between years for a given program in terms of the number of applicants.

Figure 2.9 – Histogram of the number of applications for all programs (top graphs) and by type of program (directed or non-directed, bottom graphs)



The successful applicants receive 1.37 grants on average over the relevant period.[25] The grants distribution is also skewed, but with a smaller right tail (Figure 2.10). More than 70% of the applicants are granted only once over the five years. Almost two thirds of these grants relate to the aforementioned directed programs, while the remaining one third

[24]The programs are entitled “Biology and Health”, “Ecosystems and Sustainable Development”, “Renewable Energy and Environment”, “Engineering, Methods and Security”, “Materials and Information”, “Human and Social Sciences”, and “Information and Communication Sciences and Technologies”.

[25] That is to say, 5,786 funded researchers share 7,945 grants.

Table 2.17 – Number of applications by year and by program

-	2005	2006	2007	2008	2009	Total
Non-Directed Programs	2,201	3,860	2,372	1,806	2,172	12,411
Directed Programs	3,221	4,212	2,622	1,753	1,145	12,953
Biology & Health	1,128	2,165	1,152	630	418	5,493
Ecosystems & Sustainable Development	229	202	151	145	125	852
Renewable Energy & Environment	578	447	400	304	213	1,942
Engineering, Methods & Security	71	136	202	165	66	640
Materials & Information	861	788	200	83	53	1,985
Human & Social Sciences	25	123	84	48	39	319
Information & Communication Sc. & Tech.	329	351	433	378	231	1,722
All Programs	5,422	8,072	4,994	3,559	3,317	25,364

relate to non-directed programs (Table 2.18). As for applications, directed programs are predominant among all fundings awarded at the beginning of the period, although their share decreases afterwards. However, the rise in non-directed programs over the period is less pronounced for fundings than for applications: directed and non-directed programs balance out in 2009. The number of grants is also unequal between programs, with the same features as the number of applications (see Figure 2.10).

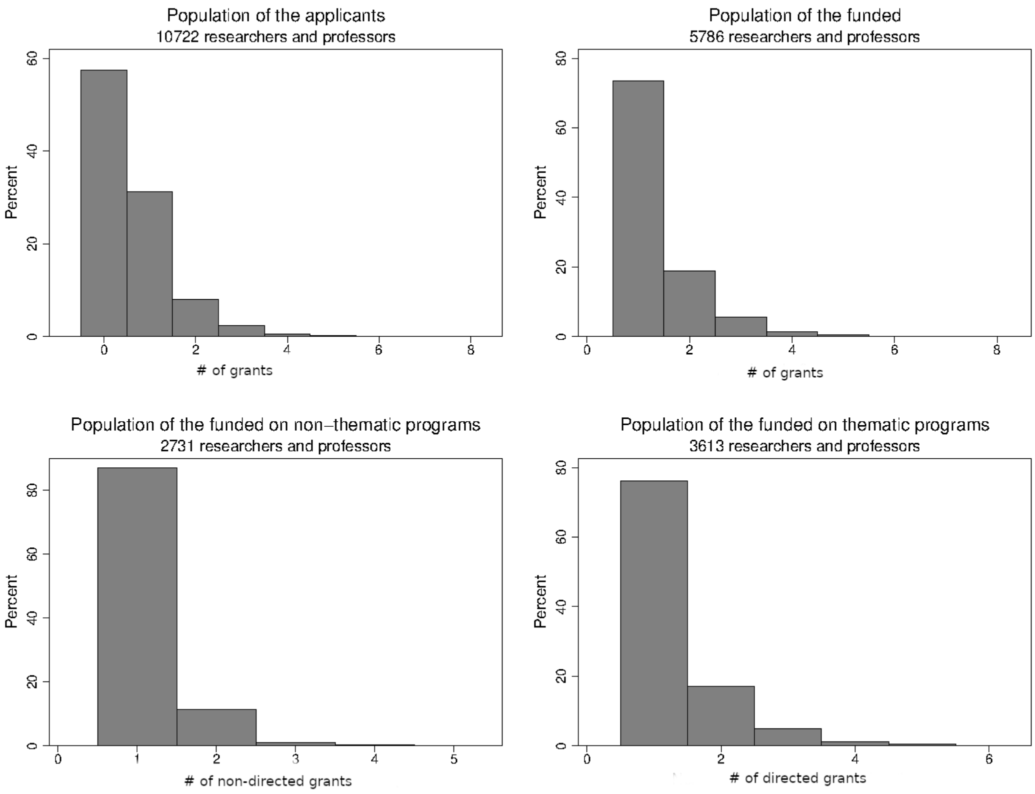
Table 2.18 – Number of grants by year and by program

-	2005	2006	2007	2008	2009	Total
Non-Directed Programs	1,001	794	581	372	385	3,133
Directed Programs	1,604	1,331	892	628	357	4,812
Biology & Health	622	576	330	170	108	1,806
Ecosystems & Sustainable Development	193	144	96	88	35	556
Renewable Energy & Environment	279	185	143	115	69	791
Engineering, Methods & Security	17	17	79	61	17	191
Materials & Information	358	279	0	0	0	637
Human & Social Sciences	3	27	22	14	3	69
Information & Communication Sc. & Tech.	132	103	222	180	125	762
All Programs	2,605	2,125	1,473	1,000	742	7,945

We now focus on the participation in ANR programs according to the scientific disciplines of the applicants. For this purpose, we first allocate sections[26] to large disciplines. This allocation turns out to be complex in a limited number of sections because of the

[26] The list of sections is given in Table 2.21. Researchers could be assigned to one of the 99 different sections, which are specific to their employing institution (if they are professors, it would be the Ministry of Research and Higher Education). The types of research centers in our database are INRA

Figure 2.10 – Histogram of the number of grants for all programs (top graphs) and by type of program (directed or non-directed, bottom graphs)



multidimensional nature of some sections. When this issue could not be resolved, allocation is made across multiple disciplines. We observe that the highest application rate is found in Physics, followed by the Life Sciences (with Chemistry and Applied Biology not far behind, see Table 2.19).

By contrast, the application rate for Mathematics is the lowest (less than half the rate for Physics). In some disciplines, such as Life Sciences, Medicine and Engineering, professors and researchers apply more frequently to directed programs, whereas non-directed programs are preferred concerning Physics, the Universe Sciences and Mathematics. In terms of the number of granted applications, the highest funding rate is found in Applied Biology and the lowest is found in Mathematics. Results by program go along with those for the applications. The prevalence of grants related to directed programs is also found in the Life Sciences, Medicine and Engineering. On the contrary, Physics, Sciences of the Universe and Mathematics are more often funded through non-directed programs. The allocations are fairly balanced between the two types of programs in Social sciences and Chemistry.

Table 2.19 – Allocation of the ANR applications into large disciplines for our final sample

	Researchers	Applicants					
		Non-Directed		Directed		Total	
Disciplines	#	#	%	#	%	#	%
Life Sciences	6,036	2,423	40	3,261	54	5,684	94
Medicine	3,478	1,055	30	1,773	51	2,828	81
Applied biology - Ecology	1,798	906	50	707	39	1,613	90
Chemistry	3,842	1,835	48	1,669	43	3,504	91
Physics	3,182	1,878	59	1,428	45	3,306	104
Sciences of the Universe	2,202	1,259	57	339	15	1,598	73
Engineering	6,441	1,845	29	3,068	48	4,913	76
Mathematics	2,646	709	27	335	13	1,044	39
Social Sciences	1,524	562	37	408	27	970	64
Total	31,149	12,472	40	12,988	42	25,460	82

Notes: The total number of researchers (respectively applicants) is 31,149 (25,460) instead of 31,081 (25,364) because of the multiple allocations of some sections to several disciplines. The number of applications in Social Sciences is low considering we excluded most Human and Social Sciences disciplines from the analysis.

In Figure 2.11 we investigate the rate of participation (number of applications and number of awards per capita) at the section level (which corresponds mainly to a sub-

(agronomic research), INRIA (computer science and engineering), INSERM (medical research), CNRS and universities, each of them has its own classification in terms of specialties.

Table 2.20 – Allocation of the ANR granted applications into large disciplines for our final sample

	Researchers	Applicants					
		Non-Directed		Directed		Total	
Disciplines	#	#	%	#	%	#	%
Life Sciences	6,036	538	09	1,132	19	1,670	28
Medicine	3,478	195	06	645	19	840	24
Applied biology - Ecology	1,798	217	12	399	22	616	34
Chemistry	3,842	428	11	508	13	936	24
Physics	3,182	541	17	480	15	1,021	32
Sciences of the Universe	2,202	341	15	156	07	497	23
Engineering	6,441	467	07	1,188	18	1,655	26
Mathematics	2,646	267	10	157	06	424	16
Social Sciences	1,524	139	09	147	10	286	19
Total	31,149	3,133	10	4,812	15	7,945	26

Notes: The number of applications in Social Sciences is low considering we excluded some Human and Social Sciences disciplines from the analysis.

discipline and an employing institution). We find a linear relationship between the rate of applications and the rate of funding, for both directed (top-right) and non-directed (top-left) programs. Some sections benefit from small positive bias in terms of the success rate (points that are on the left of the non-represented fitted straight line that could be drawn). Most are CNRS sections for non-directed programs and INSERM/INRA/INRIA sections for directed programs. When we consider the joint participation rates of sections in the two types of programs (bottom graphs of Figure 2.11), the results vary significantly depending on the sections. Some sections favor a particular type of program, while others indicate a fairly balanced participation between the directed and non-directed programs (both for applications and fundings).

Lastly, Figure 2.12 shows histograms of the size of laboratories, in terms of the number of tenured researchers or professors, in the three samples. In the majority of cases, these academics' laboratory staff is made up of 10 to 70 employees, while some of them exceed 200 tenured staff members. There is no obvious difference in size between the distributions of the three samples.

Figure 2.11 – Intensity of the participation in directed and non-directed programs at the specialties level (for sections with more than 25 researchers)

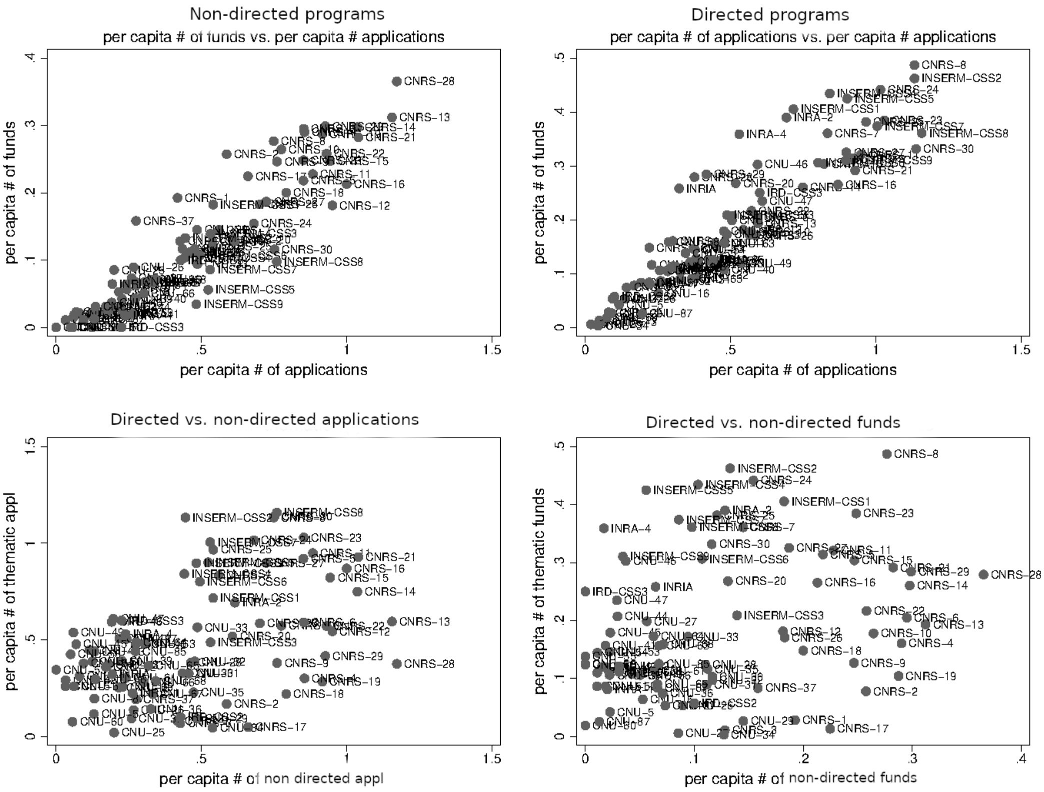


Figure 2.12 – Histogram of the size of the laboratories (number of tenured researchers or professors) in the three samples

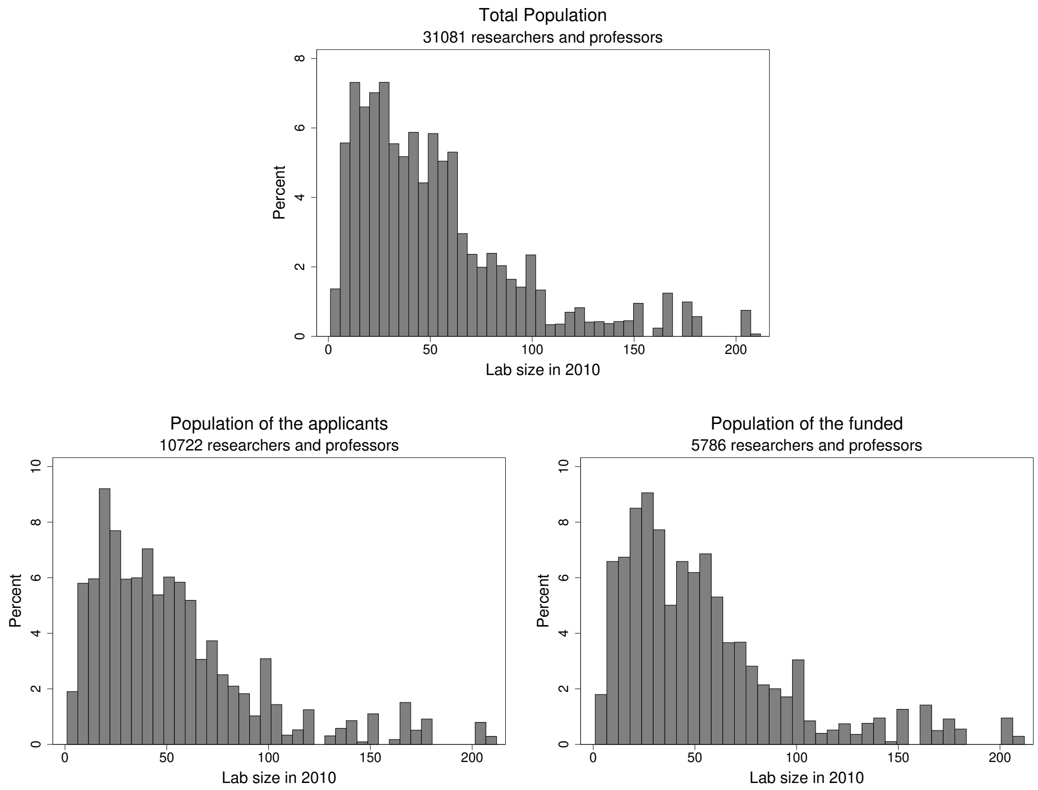


Table 2.21 – List of sections assigned to our final sample of researchers, according to the research institute

Section	
CNRS-1	Mathématiques et interactions des mathématiques
CNRS-10	Milieux fluides et réactifs : transports, transferts, procédés de transformation
CNRS-11	Systèmes supra et macromoléculaires : propriétés, fonctions, ingénierie
CNRS-12	Architectures moléculaires : synthèses, mécanismes et propriétés
CNRS-13	Physicochimie : molécules, milieux
CNRS-14	Chimie de coordination, interfaces et procédés
CNRS-15	Chimie des matériaux, nanomatériaux et procédés
CNRS-16	Chimie du vivant et pour le vivant
CNRS-17	Système solaire et univers lointain
CNRS-18	Terre et planètes telluriques : structure, histoire, modèles
CNRS-19	Système Terre : enveloppes superficielles
CNRS-2	Théories physiques : méthodes, modèles et applications
CNRS-20	Surface continentale et interfaces
CNRS-21	Bases moléculaires et structurales des fonctions du vivant
CNRS-22	Organisation, expression et évolution des génomes
CNRS-23	Biologie cellulaire : org et fonc de la cellule, pathogènes et rel hôte/pathogène
CNRS-24	Interactions cellulaires
CNRS-25	Physiologie moléculaire et intégrative
CNRS-26	Développement, évolution, reproduction, vieillissement
CNRS-27	Comportement, cognition, cerveau
CNRS-28	Biologie végétale intégrative
CNRS-29	Biodiversité, évolution et adaptations biologiques
CNRS-3	Interactions, particules, noyaux du laboratoire au cosmos
CNRS-30	Thérapeutique, médicaments et bio-ingénierie : concepts et moyens
CNRS-37	Économie et gestion
CNRS-4	Atomes et molécules, optiques et lasers, plasmas chauds
CNRS-5	Matière condensée : organisation et dynamique
CNRS-6	Matière condensée : structures et propriétés électroniques
CNRS-7	Sciences et technologies de l'information
CNRS-8	Micro et nano-technologies, élec, photo, électroma, énergie élec
CNRS-9	Ingénierie des matériaux et des structures, mécaniques de solides, acous

Table 2.21 Continued

Section	
CNU-16	Psychologie, psychologie clinique, psychologie sociale
CNU-25	Mathématiques
CNU-26	Mathématiques appliquées et applications des mathématiques
CNU-27	Informatique
CNU-28	Milieux denses et matériaux
CNU-29	Constituants élémentaires
CNU-30	Milieux dilués et optique
CNU-31	Chimie théorique, physique, analytique
CNU-32	Chimie organique, minérale, industrielle
CNU-33	Chimie des matériaux
CNU-34	Astronomie, astrophysique
CNU-35	Structure et évolution de la terre et des autres planètes
CNU-36	Terre solide : géodynamique des enveloppes supérieure, paléobiosphère
CNU-37	Météorologie, océanographie physique de l'environnement
CNU-39	Sciences physico-chimiques et technologies pharmaceutiques
CNU-40	Sciences du médicament
CNU-41	Sciences biologiques
CNU-42	Morphologie et morphogenèse
CNU-43	Biophysique et imagerie médicale
CNU-44	Biochimie, biologie cellulaire et moléculaire, physiologie et nutrition
CNU-45	Microbiologie, maladies transmissibles et hygiène
CNU-46	Santé publique, environnement et société
CNU-47	Cancérologie, génétique, hématologie, immunologie
CNU-48	Anesthésiologie, réanimation, médecine d'urgence, pharmaco et thérapeutique
CNU-49	Pathologie nerveuse et musculaire, pathologie mentale, handicap et rééducation
CNU-5	Sciences économiques
CNU-50	Pathologie ostéo-articulaire, dermatologie et chirurgie plastique
CNU-51	Pathologie cardiorespiratoire et vasculaire
CNU-52	Maladies des appareils digestif et urinaire
CNU-53	Médecine interne, gériatrie et chirurgie générale
CNU-54	Développement et pathologie de l'enfant, gynéco-obsté, endocrino et reprod
CNU-55	Pathologie de la tête et du cou
CNU-56	Développement, croissance et prévention
CNU-57	Sciences biologiques, médecine et chirurgie buccales
CNU-58	Sciences physiques et physiologiques endodontiques et prothétiques

Table 2.21 Continued

Section	
CNU-60	Mécanique, génie mécanique, génie civil
CNU-61	Génie informatique, automatique et traitement du signal
CNU-62	Energétique, génie des procédés
CNU-63	Génie électrique, électronique, photonique et systèmes
CNU-64	Biochimie et biologie moléculaire
CNU-65	Biologie cellulaire
CNU-66	Physiologie
CNU-67	Biologie des populations et écologie
CNU-68	Biologie des organismes
CNU-69	Neurosciences
CNU-85	Pharmacie en sciences physico-chimiques et ingénierie appliquée à la santé
CNU-86	Pharmacie en sciences du médicament et des autres produits de santé
CNU-87	Pharmacien sciences biologiques, fondamentales et cliniques
INRA-1	Biologie fondamentale
INRA-2	Médecine
INRA-3	Biologie/Ecologie appliquée
INRA-4	Chimie
INRA-6	Science de l'Univers
INRA-8	Mathématiques
INRA-9	Sciences humaines et sociales
INRIA	Sciences de l'ingénieur et mathématiques
INSERM-CSS1	Bases biochimiques, moléculaires et structurales du vivant
INSERM-CSS2	Génétique, épigénétique, cancérologie
INSERM-CSS3	Biologie cellulaire, développement, vieillissement
INSERM-CSS4	Physiologie et physiopathologie des syst card, vasc, pulm, néphro et musc
INSERM-CSS5	Physiologie et physiopathologie des systèmes endoc, dig, ostéo-artic et cut
INSERM-CSS6	Neurosciences, cognition, santé mentale
INSERM-CSS7	Microbiologie, immunologie, infection
INSERM-CSS8	Technologies pour la santé, thérapeutiques, biotechnologies
INSERM-CSS9	Santé publique, santé des populations : épidémio, biostat, shs
IRD-CSS1	sciences physiques et chimiques de l'environnement planétaire
IRD-CSS2	sciences biologiques et médicales
IRD-CSS3	sciences des systèmes écologiques

Appendix B. Outcome variables

In this section, we present how we built the different outcome variables used in the analysis.

Production variables

- Yearly number of contributions to articles published in WoS journals, with each paper being weighted by the inverse of the number of its authors:

$$VC_i^t = \sum_{j \in J_t} \frac{1\{i \Lleftarrow j\}}{n(j)}, \tag{2.7}$$

where J_t denotes the set of published paper in year t, $1\{.\}$ is the indicator function equal to one if the condition into brackets is verified and zero otherwise, the expression "$i \Lleftarrow j$" means i is the author of j and $n(j)$ denotes the author number of the article j. In the main paper and in the tables, we refer to this variable as the ***Volume***.

- Yearly number of articles published in WoS journals, with each paper being adjusted by the impact factor of the journal and by the inverse of the number of authors:

$$IFC_i^t = \sum_{j \in J_t} \frac{1\{i \Lleftarrow j\} \times IF(j)}{n(j)}, \tag{2.8}$$

where $IF(j)$ denotes the (three-years) impact factor of the journal where publication j has been published. In the main article and in the tables, we refer to this variable as ***Impact Factor***.

- Yearly number of articles published in WoS journals, with each article being adjusted by the number of citations in the three-year moving window $(t, t+2)$ and by the inverse of the number of coauthors:

$$CITC_i^t = \sum_{j \in J_t} \frac{\{i \Lleftarrow j\} \times C(j)}{n(j)}, \tag{2.9}$$

where $C(j)$ denotes the number of citations received by article j from articles published in the three-year moving window $(t, t+2)$. In the main body of the article and in the tables, we refer to this variable as ***Citations***.

Other outcome variables

- Mean number of authors by article in a given time period τ:

$$COA_i^\tau = \frac{\sum_{j \in J_\tau} 1\{i \Lleftarrow j\} \times n(j)}{\sum_{j \in J_\tau} 1\{i \Lleftarrow j\}}. \tag{2.10}$$

- Number of distinct coauthors recorded for the considered individual COD_i^τ.

- Number of new coauthors: number of distinct coauthors observed in period τ who did not appear among previous coauthors of i: $CODN_i^\tau$.

- Number of international collaborations: number of published articles with at least one author with a professional address located outside France for a given period τ:

$$INT_i^\tau = \sum_{j \in J_\tau} 1\{i \Lleftarrow j\} \times 1\{j \leftarrow inter\}, \tag{2.11}$$

where "$j \leftarrow inter$" means that paper j results from an international collaboration.

Appendix C. Specification of the selection model

We first discuss the basic principles used to build the selection model, before presenting the eight specifications that are firstly retained. All these specifications will be compared in the next section.

Principles

We consider two different sets of persons in which to pick controls:

- the first set consists of all the researchers and professors in our whole cleaned data set who did not get an ANR grant in the period 2005-2009, that is 25,295 persons (31,081 researchers and professors, of which 5,786 received a grant);
- the second one is a subset of the first set, which consists of researchers and professors who applied to an ANR call for proposals between 2005 and 2009, but received no funding. It comprises 4,936 persons (10,722 applicants,of which 5,786 received a grant).

Although the second group size is much smaller compared to the first one, its members are characterized by the same self-selection in terms of applying for a grant as those who were successful in doing so. Moreover, these individuals (as with the grant recipients) could have been subject to variations in their performance before applying in order to increase their chances of being selected. If researchers increase the number of authored publications before the application date, the use of the first set as control group (all non-recipients of funding) would underestimate the mean effect of funding (when using difference-in-differences method). The second group (applicants), however, has the disadvantage in terms of offering much less potential controls. Hence, non-applicants can possibly be depicted as better controls than unsuccessful grant applicants. Two types of information can be used to explain the selection process:

- Individual variables. Personal characteristics of the researchers, observed at the date of the application are likely to influence the selection of the project by the ANR, as well as future scientific production. Age is well known to affect scientific production over the career path. Scientific production first increases before eventually decreasing later in some fields. Since it is also likely to affect selection into treatment, we thus use the age of the researcher, together with the squared age to capture a possible non linear effect. We also consider several production measures built from publication data to account for scientific activity, impact and audience. We use the number of articles published in the three previous years to account for the intensity of the recent research effort, the number of citations received in the same period to control for the recent impact of the authors' research, the maximum impact factor of the journals in the same period to consider the ability to publish in large audience journals and the number of citations received over a longer period (recorded from 1999) in order to account for the long-term scientific reputation. Finally, we introduce, in some specifications, the production variation before the application year.[27] It is intended to account for the scientific production dynamics just before funding, while all the other publication variables explaining the treatment are averaged over the previous years.

- Laboratory variables. Given that the research environment quality is explicitly examined in the ANR evaluation process, laboratory attributes are likely to affect the selection of the applications (as well as the propensity to apply). They also influence the production outcomes (see for instance Carayol & Matt 2006). We select variables that measure the intensity of scientific production, the reputation in terms of citations at the laboratory level and the size of the laboratory. These variables are not included in the first specifications of the selection model because they correspond to the configuration of the laboratories in year 2010. Though mobility is limited, laboratory memberships could have changed since the application year. In theory, the model should not include covariates observed after the application date because it may bias our estimates. For instance, a grant recipient could have moved between the grant awarding year and 2010. Indeed, the recipient may now be member of a laboratory with better performance than the one he was affiliated to at the date of the application, either because the laboratory was able to employ new staff as a consequence of the grant or because the funding influenced the mobility of the recipient, which could increase the weight given to controls affiliated to laboratories with better quality. If this frequently occurs, it could result in an underestimation of the mean effect of the grant (because controls are selected in relation to better quality laboratories). However, as shown above, the inclusion of these laboratory variables does not affect our results significantly.

Some other additional relevant covariates are also considered. We use them in various forms (exact matching or explanatory covariates).

- Scientific fields. Given that the study covers scientific fields with heterogeneous publication profiles, we investigate whether the regression has to be implemented by scientific specialty. For this purpose, we investigate an exact matching with the section that also allows us to control the employer type and employment type (professor or researcher). This comes down to considering whether the conditions of selection can change from a specific section to another one. It ensures that a grant recipient will never have a control from a different field. This, however, implies a reduced set of treated and controls in each model. Some sections count a very limited

[27] These variables are calculated by taking the difference of the production measures in the level between $t-3$ and $t-1$, with t as the application date.

number of members and thus do not have enough treated or controls left to obtain consistent coefficients in the logit regression. Another disadvantage of performing an exact matching on the section variable is that the implementation turns out to be very complicated, given the large number of sections involved. Therefore, in some models, we rather perform exact matching on aggregated, thematically close sections.

- Program type. The selection process can actually follow slightly different logics according to the type of program considered. In particular, the selection processes of the directed and non-directed programs may differ. An exact matching with the program type may allow us to consider different weighting schemes of the ANR selection process, according to the type of program.

- Application year. The process for allocating ANR grants has not necessarily been the same across the years, especially in a context of the gradual establishment of the ANR. In particular, 2005, the first year of activity of the ANR, is characterized by a much higher selection rate than other years.

The selection models

We now present here the eight different specifications that we selected to estimate the propensity score. The list of variables used for model 1 to 5 are presented in Tables 2.22 to 2.24.

Model 1 The specification of the model includes individual covariates, which influence both the selection process and our outcome variables, such as the age, as well as some measures related to the scientific production in the three previous years (the number of publications adjusted for authorship, the number of citations received and the maximum impact factor of the journal). The propensity score is estimated by exact matching in terms of the section and the year of an ANR program. The control group is the whole set of French researchers who did not receive a grant from the ANR during the period 2005-2009, that is, 25,390 researchers were observed for each year. We do not consider all sections×years with less than five funded researchers. For some sections×years, the maximum likelihood algorithm of the logit regression does not converge, so we also discard these groups from the analysis. We decide not to apply any modification in the specification of the model for each group in order to avoid introducing any uncontrolled bias.

Model 2 The specification is similar to 1, with the inclusion of additional explanatory covariates related to the laboratory of each researcher.

Model 3 The control group is limited to the subset of the 4,936 unsuccessful ANR applicants.[28]We assume that this control group allows us to control the self-selection bias (the decision to apply to an ANR program). As the size of the control group is severely reduced, this specification is no longer based on exact matching on section×year. Instead, we aggregate sections on a disciplinary basis (see Table 2.25 for a description of the sections grouped together into disciplines).

Model 4 The specification is similar to 3, with the inclusion of explanatory variables related to the laboratory of the researcher.

Model 5 We assume here that the ANR selection process can be driven by various determinants, according to the type of program. Instead of using the same covariate specification for each group (formed according to section×year or large disciplines), we use a different specification of the model for directed and non-directed programs. This idea is also explained by the difficulty of finding a uniquely good specification for both program types. Compared with the previous specifications, some continuous covariates (such as the production measures) are transformed into categorical covariates. Information related to the laboratories are not included in the set of explanatory covariates. The discipline is represented by pooling sections into broad fields (see groups of sections in Table 2.26).

Models 6 to 8 These specifications are analogous to 3 to 5, with the introduction of the additional "trend" variables. We define two measures of the production evolution before the year of application. The first one refers to the variation in absolute terms of the level of production between $t-3$ and t (or $t-1$), whereas the second one refers to variation in percentage points of the output between $t-3$ and $t-1$.

$$trend\,diff = X_{t-1} - X_{t-3}$$

$$trend\,diff1 = X_t - X_{t-3}$$

$$trend\,rate = \frac{X_{t-1} - X_{t-3}}{X_{t-3}}$$

Where X denotes one of the three production resumes (volume, citations or impact Factor). The additional trend covariates used in the models 6 to 8 are:

- $trend\,diff$ where X relates to the number of citations received (in model 6)
- $trend\,rate$ where X relates to the number of citations received (in model 7)
- $trend\,diff1$ and $trend\,rate$ where X relates to the number of maximum impact factor of journal (in model 8 for non-directed programs)

[28]The control group is built using the unsuccessful ANR applicants' subset for models 3 to 8.

- $trend\ diff1$ and $trend\ rate$ where X relates first to the number of articles and then to the number of citations received (in model 8 for directed programs)

Table 2.22 – List of covariates used for the propensity score estimation in the models 1 to 4

Variable	Description	models
individual covariates	Age at the time of application	1,2
	Age squared at the time of application	1,2
	Number of publications in the previous 3 years	1,2
	Number of citations to papers published in the previous 3 years	1,2
	Maximum Impact Factor in the previous 3 years	1,2
	Total number of citations to papers published since 1999	1,2
laboratory covariates	Mean number of publications of the laboratory members in the previous 3 years	2,4
	Maximum Impact Factor in the previous 3 years of the top member of the lab	2,4
	Size of the lab	2,4
covariates used for exact matching	Section related to the classification of the research institute at which the researcher is affiliated	1,2
	Year of the ANR selection of project recipients	1,2
	Relatively large Scientific disciplines dummies (cf Table 2.25)	3,4

Notes: All outcome variables are adjusted for co-authorship (fractional counting). All quantitative variables are in levels, apart from two of these variables categorized in four classes (4 dummies are created): top 10%, next 20%, next 30% and last 40% for the Total number of citations and in quartiles for the Size of the lab.

Table 2.23 – List of covariates used for the propensity score estimation in the model 5 (non-directed programs)

Non-directed programs
Variable Description
Age at the time of application
Number of publications in the previous 3 years
Number of citations to papers published in the previous 3 years
Maximum Impact Factor in the previous 3 years
Total number of citations to papers published since 1999
Large Scientific disciplines dummies (see Table 2.26)
Dummies when an university or a specific research institute is the employer

Notes: All outcome variables are adjusted for co-authorship (fractional counting) and categorized in four classes (four dummies are created): top 10%, next 20%, next 30% and last 40%. The large scientific disciplines dummies are Life sciences, Medicine, Chemistry, Physics, the Universe Sciences, Engineering, Mathematics, Information science, Human & social sciences. The specific research institutes are: CNRS, INRA, INRIA, IRD, and INSERM.

Table 2.24 – List of covariates used for the propensity score estimation in the model 5 (for directed programs)

directed programs
Description
age at the time of application
Number of publications in the previous 3 years
Number of citations to papers published in the previous 3 years
Maximum Impact Factor in the previous 3 years
Total number of citations to papers published since 1999
The specific directed program
Year of the application
Interaction between the specific directed program theme and the application year

Notes: All outcome variables are adjusted for co-authorship (fractional counting) and categorized in four classes (four dummies are created): top 10%, next 20%, next 30% and last 40%.

Table 2.25 – Groups of sections , given the classification of the research institute (used in model 3 and model 4)

Groups of sections (by research institute)
CNU-25 -26
CNRS-11 -12 -13
CNU-37 -35 -36
CNRS-23 -20 -21
CNRS-26 -25 -27 -24 -28
CNU-68 -65 -66 -41
INSERM-CSS8 -CSS7
CNU-40 -39
CNU-52 -43 -45 -57 -56 -50 -46 -44 -53 -49 -51 -54 -42 -47 -55 -48 -58
INSERM-CSS1 -CSS3 -CSS6 -CSS5
CNRS-38 -31
CNRS-4 -2 -3
CNU-29 -30
CNRS-40 -36
CNU-4 -3 -1 -2
CNU-5 -6
CNRS-39 -31
CNU-23 -24
CNU-7 -71
CNU-73 -13 -14 -15 -10 -8 -12 -9 -11
CNU-76 -18 -17 -72 -77
CNRS-5 -6
CNRS-17 -15
CNRS-18 -16
CNU-27 -61

Table 2.26 – Groups of sections (used in model 5 and model 8)

Discipline	Groups of sections
Life sciences	CNRS-20 -21 -22 -23 -24 -25 -26 -27 -28 -29 -30 ; CNU-39 -40 -41 CNU-64 -65 -66 -67 -68 -69 INRA-1 -3 INSERM-CSS2 -CSS4 -CSS7 -CSS8
Medical research	CNU-42 -43 -44 -45 -46 -47 -48 -49 -50 -51 -52 -53 -54 -55 -57 -58 -85 -86 -87 INRA-2 INSERM-CSS1 -CSS3 -CSS5 -CSS6
Chemistry	CNRS-15 -16 -17 -18 -19 CNU-31 -32 -33 INRA-4
Physics	CNRS-2 -3 -4 -5 -6 CNU-28 -29 -30
Universe sciences	CNRS-11 -12 -13 -14 CNU-34 -35 -36 -37 INRA-6
Engineering	CNRS-10 -9 CNU-60 -62 INRIA
Mathematics	CNRS-1 CNU-25 -26 INRA-8
ICST	CNRS-7 -8 CNU-27 -61 -63
Human & social sciences	CNRS-37 CNU-16 -5 INRA-9 INSERM-CSS9
Others (IRD)	IRD-CSS1 -CSS2 -CSS3

Appendix D. A parallel path test before treatment

The conditional difference-in-difference model is valid if the parallel trend assumption is verified. It states that the outcome variable for the treated should have experienced (after the treatment date) the same progress on average, in the absence of treatment, as the controls who have the same probability of assignment into treatment $p(x)$. It can be written as follows:

$$E(Y_{t+\tau} - Y_t | T = 1, P(X)) = E(Y_{t+\tau} - Y_t | T = 0, P(X)), \tag{2.12}$$

where Y is the outcome variable observed at the year of application t and, in a later year, at $t+\tau$, while T denotes the decision of the ANR to select the project or not and $P(X)$ is the propensity score. The parallel path assumption in Equation 2.12, however, cannot be tested directly because the counterfactual outcome of the treated is not available. That said, we can compare the outcome paths of the treated and the controls before treatment. We set up a parallel path test on the period before the attribution of the treatment. We assume that individuals who follow parallel trajectories right before the assignment are also likely to share parallel paths afterwards (all other factors being equal). Our objective is to check whether the production difference between $t-3$ and t is significantly different (in weighted means) between the treated and the controls for each specification of the selection model and for each matching method. The test is based on a difference-in-differences model before the application year. We want to check whether the variation in outcomes during the three years before the selection of grant recipients (between $t-3$ and $t-1$) is significantly different between the controls and those who received grants. If the results show a significant difference, it would disprove our assumption of a parallel trend between controls and treated. The results are presented in Table 2.27. A robustness check is presented in Table 2.28 in which we compare outcomes between $t-3$ with t.

The main results of the parallel path tests between $t-3$ with $t-1$ are the following:

- Only Models 1 and 2 do not pass the tests (a significant difference of trajectories between treated and controls).[29]
- Models 3 to 8 exhibit very weak and insignificant differences in the production dynamics between groups for the three outcome measures.[30]

We then repeat the tests by comparing the outcomes of year $t-3$ with the outcomes of year t.

- Only the model 8 returns non-significant differences between groups, whatever the technique used to form the control group.

[29] The difference is also significant for the models 3 and 4 of the IPTW approach, when the citations measure is the outcome.

[30] Only the models 3 and 4 with the IPTW method exhibit significant differences.

- The model 5 also passes the test when implemented using the five nearest neighbors technique.

Although it is complicated to order the different models according to the quality of the results obtained from the test, we can assert that the model 8 provides the most relevant estimation, as, for any of its weighting schemes and for any one of the three outcome measures, the parallel path hypothesis before treatment is never violated.

Table 2.27 – Parallel path test : Difference-in-differences estimates of the mean effect of treatment on various production variables (calculated from $t-3$ to $t-1$)

	δ^{5nn}		δ^{kernel}		δ^{iptw}	
	1	2	1	2	1	2
Volume	.00262	.00592	.00827	.00770	.01265**	.01098**
	(0.48)	(1.03)	(1.62)	(1.44)	(2.34)	(2.05)
Citations	.01915	.02959**	.02533**	.03738***	.03446***	.03269***
	(1.55)	(2.36)	(2.18)	(3.17)	(2.86)	(2.78)
Impact Factor	.01941**	.01777.**	.01788**	.02075**	.01665**	.01754**
	(2.31)	(2.02)	(2.27)	(2.52)	(2.16)	(2.18)

	δ^{5nn}			δ^{kernel}			δ^{iptw}		
	3	4	5	3	4	5	3	4	5
Volume	.00333	.00202	-.00650	.00533	.00353	-.00700	.00723	.00969	-.00518
	(0.48)	(0.29)	(-0.95)	(0.83)	(0.53)	(-1.1)	(1.06)	(1.17)	(-0.82)
Citations	.02358	.01399	-.00451	.01862	.0137	.00291	.02546*	.02523*	.00233
	(1.57)	(0.93)	(-0.30)	(1.31)	(0.92)	(0.21)	(1.73)	(1.68)	(0.17)
Impact Factor	.00781	-.00131	-.00571	.00496	-.00257	-.00192	.00595	.00106	-.00627
	(0.77)	(-0.13)	(-0.65)	(0.51)	(-0.26)	(-0.24)	(0.61)	(0.10)	(-0.78)

	δ^{5nn}			δ^{kernel}			δ^{iptw}		
	6	7	8	6	7	8	6	7	8
Volume	.00074	.00442	-.00954	.00108	.00472	-.00809	.00372	.00742	-.00727
	(0.11)	(0.65)	(-1.39)	(0.17)	(0.72)	(-1.27)	(0.53)	(0.87)	(-1.13)
Citations	.01295	.00322	-.00454	.00719	.00328	-.0012	.0088	.01153	-.00268
	(0.86)	(0.21)	(-0.3)	(0.5)	(0.22)	(-0.09)	(0.58)	(0.75)	(-0.19)
Impact Factor	.00184	-.00989	-.00515	.00319	-.00925	-.00383	.00224	-.00148	-.00588
	(0.18)	(-0.96)	(-0.6)	(0.33)	(-0.94)	(-0.47)	(0.23)	(-0.14)	(-0.72)

*Notes: Conditional difference-in-difference results. Dependent variables in Log. Robust t-stats in parentheses, clustered at the project level. Significance levels: 0.01: ***, 0.05: **, 0.10: *.*

Table 2.28 – Parallel path test : Difference-in-differences estimates of the mean effect of treatment on various production variables (calculated from $t-3$ to t)

	δ^{5nn}		δ^{kernel}		δ^{iptw}	
	1	2	1	2	1	2
Volume	-.03863***	-.03716***	-.05787***	-.05582***	-.04238***	-.042***
	(-5.4)	(-5.01)	(-8.79)	(-8.13)	(-6.54)	(-6.33)
Citations	-.11242***	-.11238***	-.1183***	-.12586***	-.1108***	-.11078***
	(-7.64)	(-7.42)	(-8.55)	(-8.89)	(-8.09)	(-7.84)
Impact Factor	-.07165***	-.07123***	-.085***	-.08648***	-.0697***	-.06944***
	(-6.91)	(-6.65)	(-9.01)	(-8.65)	(-7.81)	(-7.28)

	δ^{5nn}			δ^{kernel}			δ^{iptw}		
	3	4	5	3	4	5	3	4	5
Volume	-.01788**	-.00986	-.00336	-.0145*	-.01339	-.00207	-.01615*	-.00941	-.00320
	(-1.97)	(-1.02)	(-0.47)	(-1.7)	(-1.43)	(-0.32)	(-1.87)	(-1.06)	(-0.49)
Citations	-.05572***	-.03816*	-.02149	-.04503***	-.04739**	-.02437*	-.04416**	-.04537*	-.02464*
	(-3.03)	(-1.99)	(-1.4)	(-2.62)	(-2.54)	(-1.75)	(-2.54)	(-2.5)	(-1.78)
Impact Factor	-.02595**	-.01843	-.00817	-.02275*	-.02203	-.00657	-.02197*	-.00984	-.00492
	(-2.01)	(-1.38)	(-0.88)	(-1.88)	(-1.72)	(-0.78)	(-1.81)	(-0.69)	(-0.57)

	δ^{5nn}			δ^{kernel}			δ^{iptw}		
	6	7	8	6	7	8	6	7	8
Volume	-.01875**	-.01241	.00078	-.01363	-.0139	-.00101	-.01202	-.00751	-.00139
	(-2.06)	(-1.31)	(0.11)	(-1.61)	(-1.54)	(-0.15)	(-1.34)	(-0.85)	(-0.21)
Citations	-.04165**	-.03541*	-.01868	-.03777**	-.03951**	-.02018	-.03229*	-.03657**	-.02246
	(-2.24)	(-1.85)	(-1.23)	(-2.16)	(-2.13)	(-1.45)	(-1.8)	(-1.98)	(-1.58)
Impact Factor	-.02285*	-.01457	-.00626	-.02355*	-.01771	-.004.	-.02059*	-.00773	-.00484
	(-1.73)	(-1.1)	(-0.69)	(-1.94)	(-1.4)	(-0.47)	(-1.67)	(-0.56)	(-0.56)

*Notes: Conditional difference-in-difference results. Dependent variables in Log. Robust t-stats in parentheses, clustered at the project level, clustered at the project level. Significance levels: 0.01: ***, 0.05: **, 0.10: *.*

Appendix E. Balance diagnostics

We present the balance tests applied to the eight specifications of our selection model. Following Austin (2011), such test builds upon the idea that, if the CIA assumption holds, treated and controls should share a similar distribution of their observables X after matching/weighting. That is to say, for each level of the estimated propensity score, the distribution of the attributes X is conditionally independent of treatment status. If this balancing property is satisfied, i.e. covariates in X are balanced between treated and control subgroups for all propensity scores, then we can reliably assume that the conditional assignment into treatment is random. The difficulty to find several treated and controls for each level of $p(X)$ makes this assumption not testable in this way. Instead, we first implement a balance test without any restriction on the value of the propensity score. We next refine the test by dividing the range of the propensity score into several intervals.

For each specification of the model and for each weighting technique used, we calculate the standardized difference (in %), which stands for the remaining bias between groups (Austin 2009), a follows:

$$standt.bias = \frac{\bar{x}_{T=1} - \bar{x}_{T=0}}{\sqrt{\left(s^2_{T=1} + s^2_{T=0}\right)/2}} \times 100, \tag{2.13}$$

where $\bar{x}$ and s^2 respectively denote the weighted mean and variance of the covariates among the treated $(T = 1)$ and the controls $(T = 0)$.[31]

Figures 2.13 to 2.16 report the distribution of the estimated propensity score for the directed and the non-directed programs for our preferred model (model 8), with the nearest neighbors and the kernel matching approaches (Tables associated with the IPTW scheme are presented in chapter 2). We observe that controls tend to have a lower (estimated) probability to be treated compared to the grantees. After weighting, the propensity score is similarly distributed between treated and controls, as well as what we found with the IPTW weighting methods, but the distribution are even more close with the nearest neighbors approach.

Then, we present the balance test for each covariate used in the selection model in Figures 2.17 and 2.18. The standardized differences are reported for each covariate in line, for directed and non-directed programs and for the two matching methods. We see that the existing bias between groups has been severely reduced after matching, while the standardized difference is low for each specification (far below the 11% threshold usually retained in the literature). All the other specifications we present in this paper satisfy such balancing properties.[32]

[31]Equation 2.13 is used to calculate the standardized mean for a continuous variable. The calculation is slightly different when we refer to a categorical variable.

[32]Other balance tests are not presented due to space constraint.

We next refine the balance test in terms of dividing the range of the estimated propensity scores into several strata where the balancing property holds (Rosenbaum & Rubin 1984, Dehejia & Wahba 1999, Austin 2009). Following the algorithm used in Dehejia & Wahba (2002) and Becker and Ichino (2002) , we proceed as follow:

- A limited number of intervals is chosen so that we find an equal mean value of the propensity score for the treated and control subsamples.

- We implement the covariates balance test in each previously defined stratum of $p(x)$. If the equality of the means of a covariate between the treated and control subsamples does not hold, we reduce the size of the interval or finally change the specification of the model (in introducing interaction terms, for example).

For each of the directed and non-directed specifications of the model 8, we are able to divide the propensity scale into seven strata, into which all covariates are balanced.

Figure 2.13 – Density and box plot of the estimated propensity scores before and after matching with the 5 nearest neighbors method for the directed programs

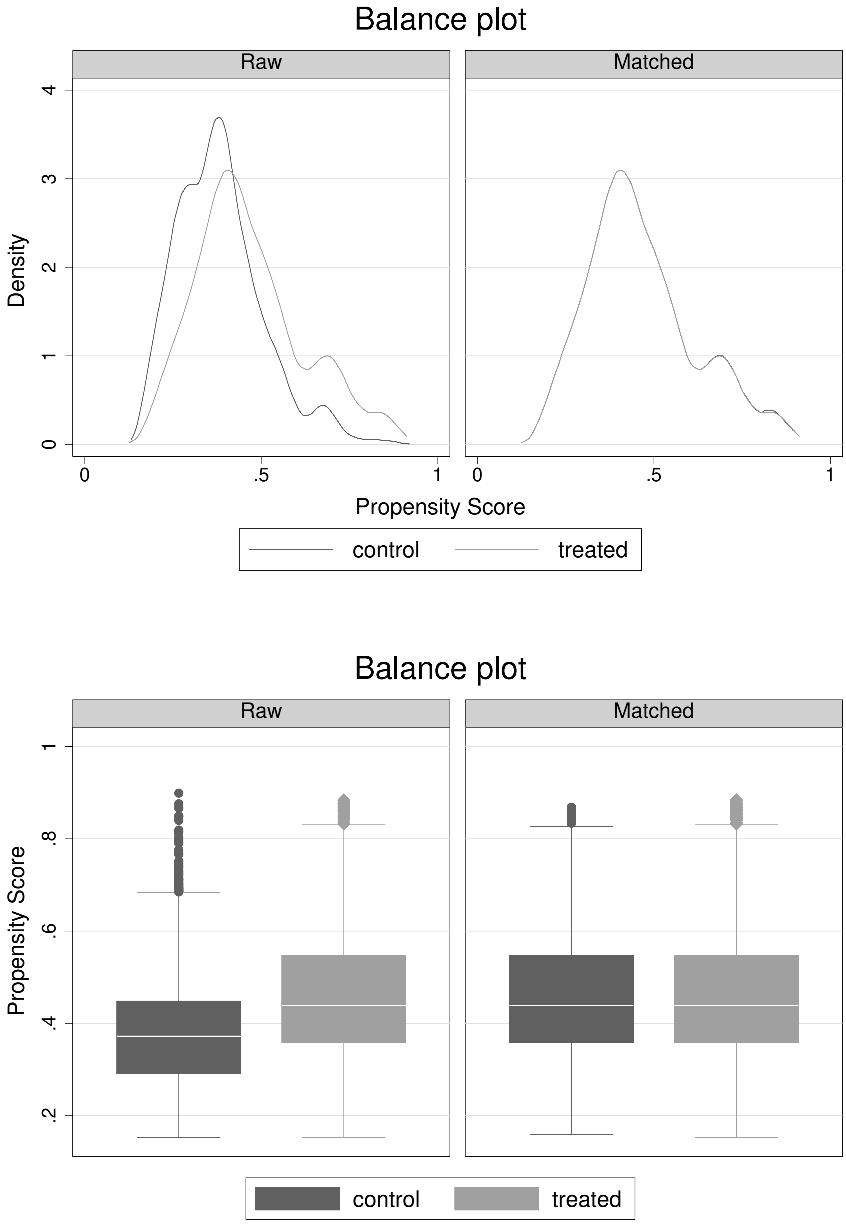


Figure 2.14 – Density and box plot of the estimated propensity scores before and after matching with the kernel method for the directed programs

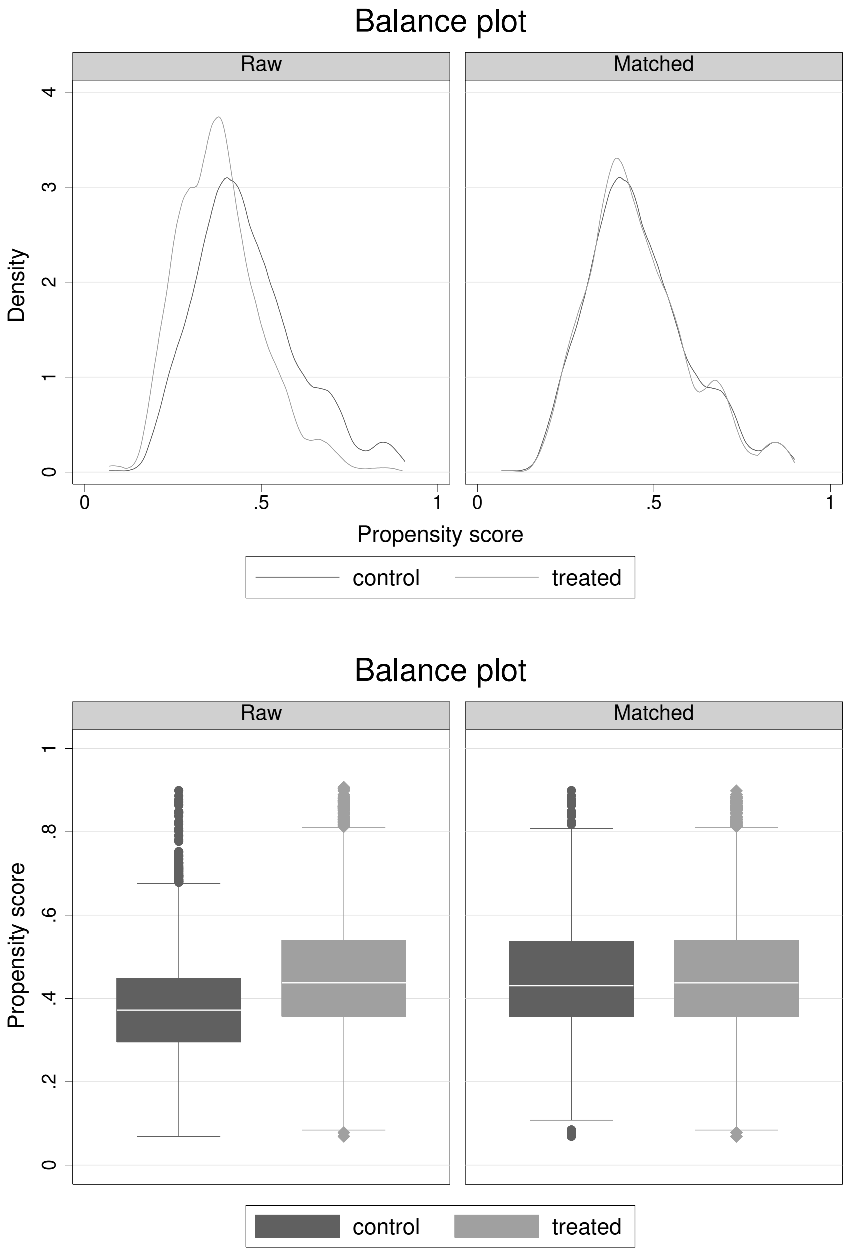


Figure 2.15 – Density and box plot of the estimated propensity scores before and after matching with the 5 nearest neighbors method for the non-directed programs

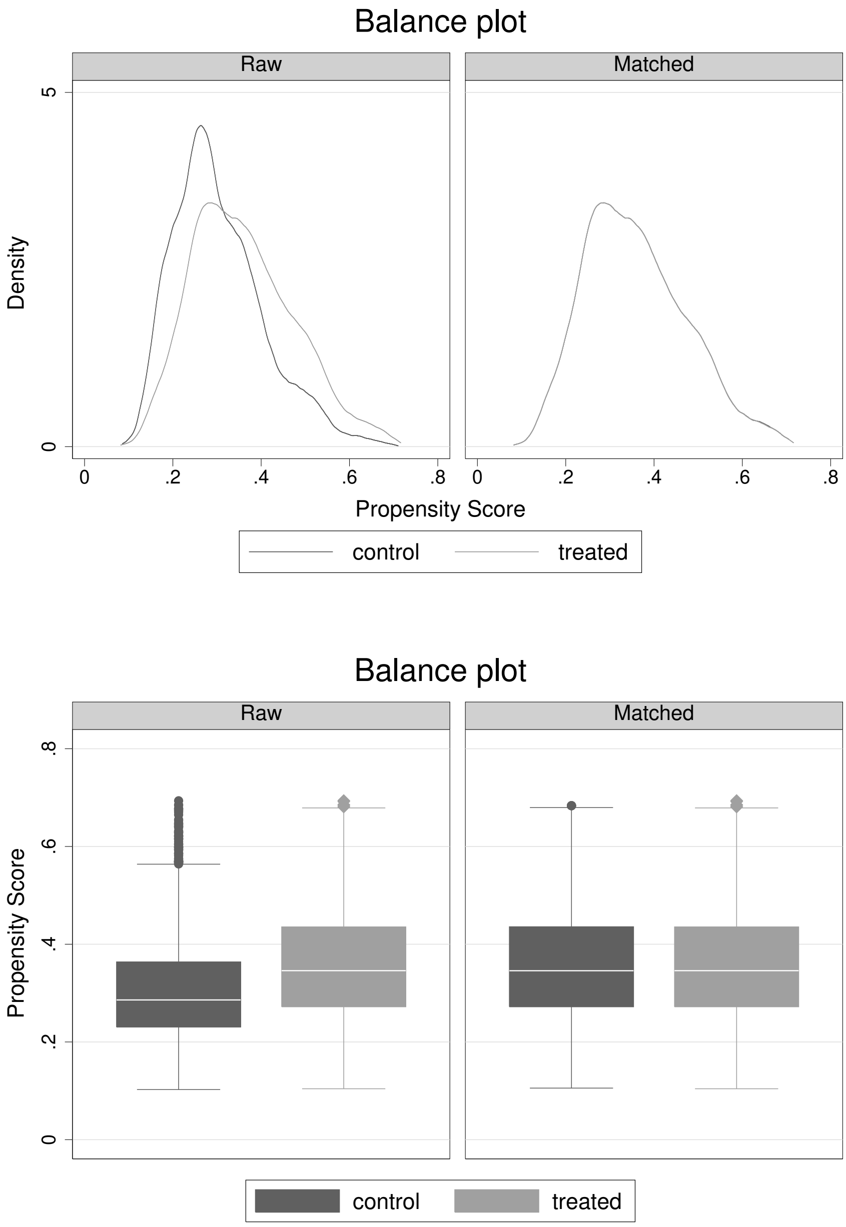
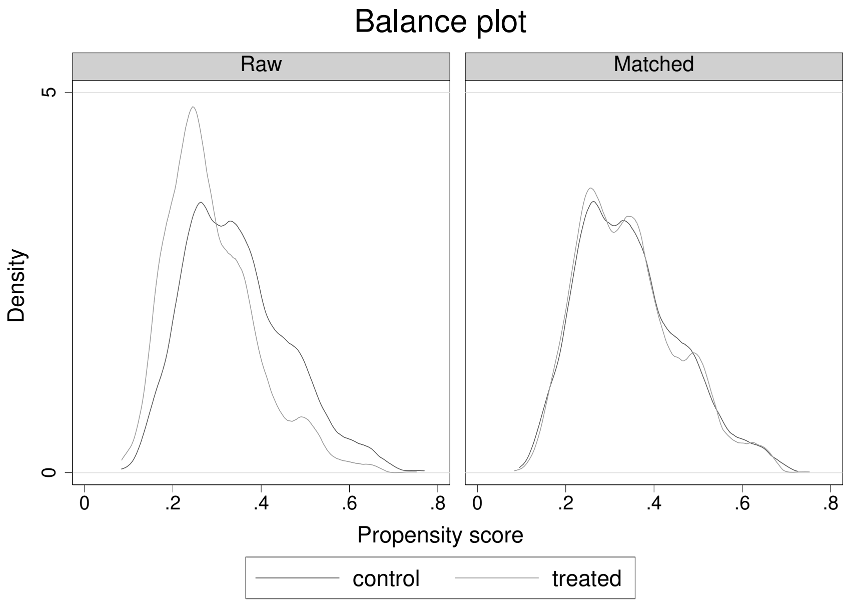


Figure 2.16 – Density and box plot of the estimated propensity scores before and after matching with the kernel method for the non-directed programs



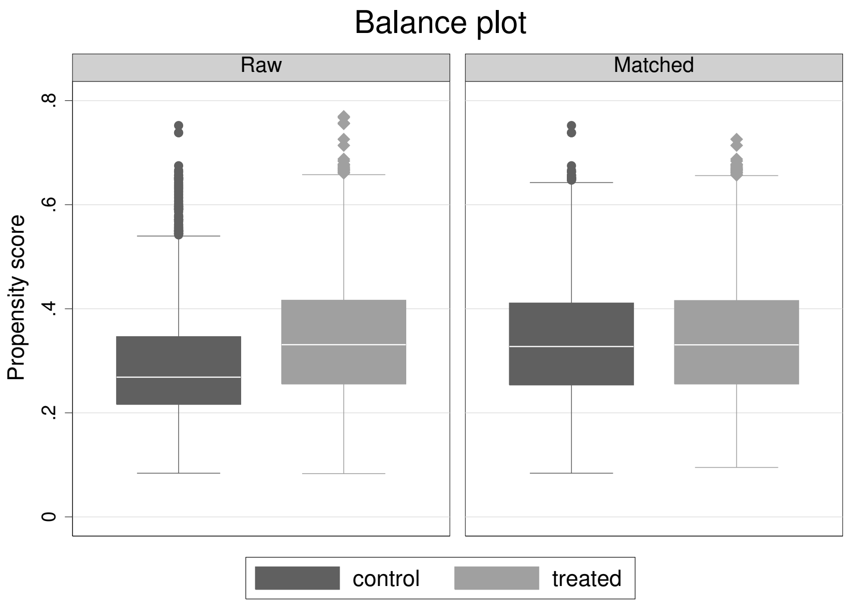
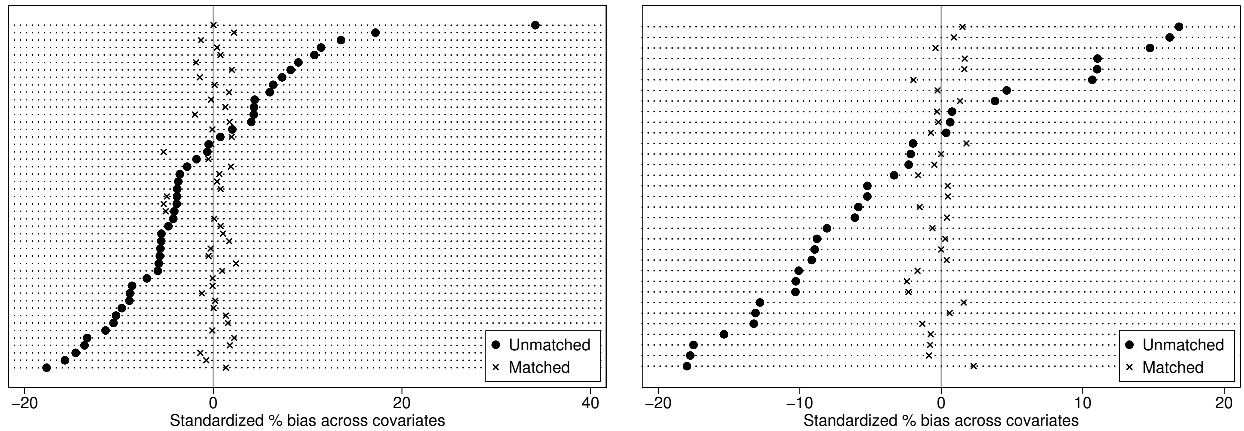
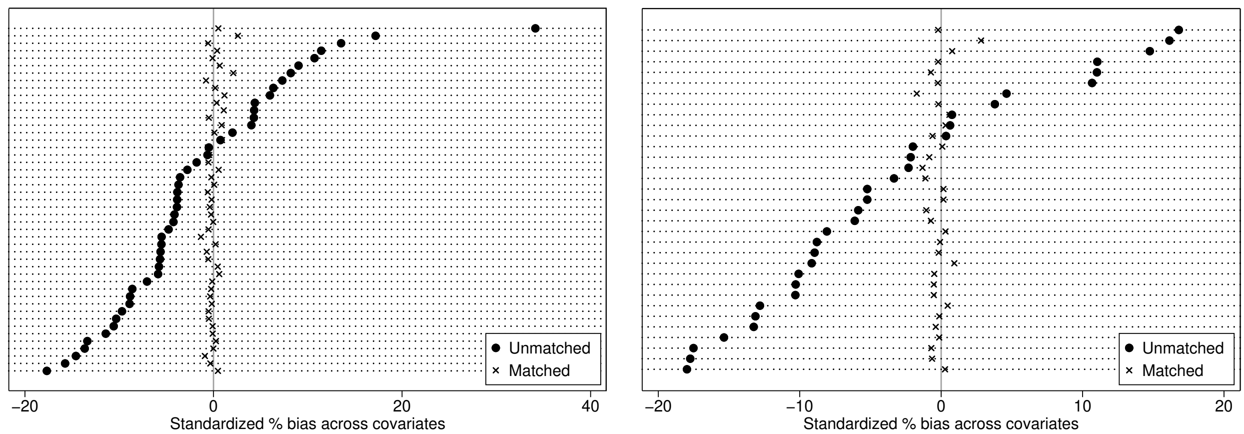


Figure 2.17 – Standardized bias (in %) associated with each explanatory covariates in the original unmatched sample and in the matched sample for the directed (left graph) and non-directed (right graph) programs with the nearest neighbors method



Notes: Each dotted line represents an explanatory covariate included in the X vector.

Figure 2.18 – Standardized bias (in %) associated with each explanatory covariates in the original unmatched sample and in the matched sample for the directed (left graph) and non-directed (right graph) programs with the kernel method



Notes: Each dotted line represents an explanatory covariate included in the X vector.

Appendix F. Supplementary estimation results

We present the results associates with our preferred model (model 8) and according to the matching or weighting approaches used to build the control group. We find that the choice of the matching/weighting method weakly influences the magnitude of our results. Only in a few cases, we observe a small difference with the nearest neighbor matching approach, which can be explained by the reduced size of the constructed sample.

Table 2.29 – Average treatment effect of receiving an ANR grant on outcomes (the three years after treatment against the three years before).

	δ^{5nn}	δ^{kernel}	δ^{iptw}
Volume	.03738***	.03544***	.03503***
	(4.46)	(4.54)	(4.46)
Citations	.1428***	.15098***	.15254***
	(8.42)	(9.45)	(9.30)
Impact Factor	.08023***	.08206***	.08252***
	(7.0)	(7.64)	(7.53)
Average	0.0201**	0.0217***	0.0218***
Team Size	(2.37)	(2.76)	(2.71)
Coauthors	0.0930***	0.0984***	0.0981***
	(6.24)	(7.09)	(7.02)
International	0.0437***	0.0414***	0.0418***
Collaborations	(2.74)	(2.81)	(2.82)
New Coauthors[a]	0.0595**	0.0651***	0.0668***
	(2.54)	(2.97)	(3.03)

*Notes: Conditional difference-in-difference results. Coefficients and standard errors of the interaction term between the post-funding period dummy and the treatment dummy in a fixed effect regression. Observations are weighted either according to the nearest neighbors, to the kernel or to the inverse probability of treatment. Dependent variables in Log. Robust t-stats in parentheses, clustered at the project level, clustered at the project level. Significance levels: 0.01: ***, 0.05: **, 0.10: *.*

Table 2.30 – Differentiated effects of receiving an ANR grant on outcomes according to non-directed versus directed funding schemes (the three years after treatment against the three years before).

	δ^{5nn}	δ^{kernel}	δ^{iptw}
Volume	0.02136	0.02858*	0.02777*
	(1.28)	(1.85)	(1.77)
Citations	0.1654***	0.2012***	0.2028***
	(4.92)	(6.37)	(6.26)
Impact Factor	0.0899***	0.1128***	0.1111***
	(3.94)	(5.28)	(5.10)
Average	0.0087	0.0004	-0.0011
Team Size	(0.51)	(0.03)	(-0.07)
Coauthors	0.0247	0.0199	0.0201
	(0.83)	(0.72)	(0.72)
International	0.0270	0.0282	0.0286
Collaborations	(0.85)	(0.96)	(0.97)
New Coauthors[a]	-0.047	-0.05	-0.049
	(-0.95)	(-1.02)	(-1.12)

*Notes: Conditional difference-in-difference-in-difference results. Coefficients and standard errors of the triple interaction term between the post-funding period dummy, the treatment dummy and the non-directed-program dummy, in a fixed effect regression. Observations are weighted either according to the nearest neighbors, to the kernel or to the inverse probability of treatment.Dependent variables in Log. Robust t-stats in parentheses, clustered at the project level. Significance levels: 0.01: ***, 0.05: **, 0.10: *.*

[a] *Conditional differences results only as this variable counts the new items in the post-treatment period as compared to the pre-treatment period.*

Table 2.31 – Differentiated effects of receiving an ANR grant on outcomes according to age dummy: below the median age (43) versus over (the three years after treatment against the three years before).

	δ^{5nn}	δ^{kernel}	δ^{iptw}
volume	0.0253	0.0202	0.0221
	(1.51)	(1.30)	(1.41)
Citations	0.107***	0.0866***	0.0952***
	(3.32)	(2.86)	(3.09)
Impact Factor	0.0359	0.0228	0.0266
	(1.62)	(1.11)	(1.29)
Average	-0.0338**	-0.0251*	-0.0279*
Team Size	(-2.05)	(-1.66)	(-1.80)
Coauthors	0.008	-0.0047	-0.0034
	(0.27)	(-0.17)	(-0.13)
International	0.0421	0.287	0.0280
Collaborations	(1.32)	(0.98)	(0.95)
New Coauthors[a]	0.0035	-0.0416	-0.0568
	(0.07)	(-0.86)	(-0.83)

*Notes: Conditional difference-in-difference-in-difference results. Coefficients and standard errors of the triple interaction term between the post-funding period dummy, the treatment dummy and the below-the-median-age dummy, in a fixed effect regression. Observations are weighted either according to the nearest neighbors, to the kernel or to the inverse probability of treatment.Dependent variables in Log. Robust t-stats in parentheses, clustered at the project level. Significance levels: 0.01: ***, 0.05: **, 0.10: *.*

[a] *Conditional difference-in-differences results only as this variable counts the new items in the post-treatment period as compared to the pre-treatment period.*

Table 2.32 – Differentiated effects of receiving an ANR grant on publication outcomes (next three years against previous three years) according to the investigator's role (principal investigator vs. partner coordinator).

	δ^{5nn}	δ^{kernel}	δ^{iptw}
volume	-0.0235	-0.0147	-0.0160
	(-1.37)	(-0.91)	(-0.99)
Citations	0.00602	0.0114	0.00500
	(0.18)	(0.36)	(0.15)
Impact Factor	0.00847	0.0219	0.0179
	(0.37)	(1.02)	(0.82)

*Notes: Conditional difference-in-difference-in-difference results. Coefficients and standard errors of the triple interaction term between the post-funding period dummy, the treatment dummy and the project-principal-investigator (PI) dummy, in a fixed effect regression. Observations are weighted either according to the nearest neighbors, to the kernel or to the inverse probability of treatment.Dependent variables in Log. Robust t-stats in parentheses, clustered at the project level, clustered at the project level. Significance levels: 0.01: ***, 0.05: **, 0.10: *.*

Table 2.33 – Differentiated effects of receiving an ANR grant on publication outcomes according to the year of funding, on the main production variables (next three years against previous three years).

Year	Volume	Impact Factor	Citations
2006	0.0019	0.0108	-0.0099
	(0.06)	(0.26)	(-0.16)
2007	0.0169	0.047	-0.0065
	(0.52)	(1.09)	(-0.10)
2008	-0.0146	0.0317	-0.025
	(-0.43)	(0.72)	(-0.38)
2009	0.0322	0.0303	0.0064
	(0.97)	(0.72)	(0.10)

*Notes: Conditional difference-in-difference-in-difference results. Coefficients and standard errors of the triple interaction term between the post-funding period dummy, the treatment dummy and the year considered (the year 2005 is in reference), in a fixed effect regression. Observations are weighted according to the inverse probability of treatment. Dependent variables in Log. Robust t-stats in parentheses, clustered at the project level. Significance levels: 0.01: ***, 0.05: **, 0.10: *.*

Table 2.34 – Differentiated effects of receiving an ANR grant on publication outcomes according to the scientific discipline of the applicant (the three years after treatment against the three years before).

Field of science	Volume	Impact Factor	Citations
Medicine	-0.0266	-0.0624	-0.0366
	(-0.96)	(-1.48)	(-0.61)
Chemistry	-0.0153	-0.0244	0.0176
	(-0.61)	(-0.64)	(0.32)
Physics	0.0362	0.00717	0.0934
	(1.31)	(0.17)	(1.57)
Engineering	0.0190	-0.0266	0.0690
	(0.61)	(-0.61)	(1.02)
Universe Sciences	0.0167	0.0273	0.103
	(0.57)	(0.60)	(1.60)
ICST[a]	0.0616**	0.0120	0.0881*
	(2.50)	(0.37)	(1.74)
Mathematics	0.00617	-0.0213	0.0743
	(0.13)	(-0.38)	(0.84)
Social Sciences	-0.0203	-0.0388	0.00250
	(-0.41)	(-0.71)	(0.03)

*Notes: Conditional difference-in-difference-in-difference results. Coefficients and standard errors of the triple interaction term between the post-funding period dummy, the treatment dummy and the scientific discipline of the applicant (the Life sciences are in reference), in a fixed effect regression. Observations are weighted according to the inverse probability of treatment. Dependent variables in Log. Robust t-stats in parentheses, clustered at the project level. Significance levels: 0.01: ***, 0.05: **, 0.10: *.*

[a] *ICST refers to "Information and Communication Sciences and Technologies".*

Appendix G. Authors disambiguation

The disambiguation algorithm

In this section, we present the three main stages of our disambiguation algorithm, as well as some descriptive elements about the implementation of the procedure. To be definitely selected, a document has to pass the seed stage or, if not, the expand stage. We now present these two stages.

The seed stage

The seed stage consists of four conditions that need to be jointly verified:

- The surname and the first name initial(s) of the researcher should be identified within author identities (presented with a surname and first name initials). The matching allows for minor variations in the name, such as some additional first name initials.
- The publication date of the article should be consistent with the researchers' age that year. We have retained a minimal age of 24 years old and an upper limit of 80 years old.
- The fine-grained scientific field of the researcher (defined as the section) should be consistent with the specialty associated with the journals which publish the papers. For this purpose, we use the classification of scientific journals into disciplines performed by the *Obvervatoire des Sciences et Techniques*.
- The researcher's institution should be mentioned in the addresses of the authors' affiliations . In order to be able to establish a connection between both information types (they could be spelled differently), the complete denomination of the institution is chosen (e.g., Université d'Aix-Marseille, Université Toulouse III or ENS Paris). Moreover, since some universities have merged since 2000, we consider both the former and the recent names of the university in this case. (e.g., Université de Bordeaux). We did not use the laboratory name to perform the comparison because of a larger variation in the spelling options, and to allow the affiliation of a researcher to more than one research laboratory in the same institution.

The expand stage

The expand stage offers a chance to all the documents that did not pass the seed stage to be validated. We relax some of the previous conditions, and substitute them by new requirements which are based on the potential similarity between papers published by a researcher. We consider three types of information:

- Two types of keywords (those reported by the authors and those attributed by ISI WoS),
- The coauthors (surname and first name initials),
- The reference lists.

The basic idea is that scientists are more likely to use the same keywords, work with the same people and cite the same papers. In each step, we focus on the similarity between the sets of validated and not validated papers for each of the researchers.

Basically, the expand stage works as follows:

1. First, relax the fourth condition of the seed stage (same institution);[33]
2. Then validate all candidate papers reporting at least one keyword (reported by the authors) already assigned to a validated article;
3. If some articles are validated in step 2, add them to the list of previously selected article and return to 2; otherwise, go to next step;
4. Then validate all candidate papers reporting at least one keyword (attributed by ISI WoS) already assigned to validated articles;
5. If some articles are validated in step 4, add them to the list of previously selected article and return to 4; otherwise, go to next step;
6. Then validate all candidate papers that are authored by at least one of the authors of the articles previously validated;[34]
7. If there are validated articles in step 6, add them to the list of previously selected article and return to 6; otherwise, go to next step;
8. If no article is validated in steps 2, 4 and 6, go to stage 9; otherwise, loop on step 2;
9. Now relax the third condition of the seed stage (same field);[35]
10. Then validate all candidate papers that have a reference list sufficiently similar to at least one of the articles previously validated;
11. If some articles are validated in step 10, add them to the list of previously selected article and return to 10; then stop anyway after 30 loops;[36]

[33] This strategy considers that scientists can be mobile and thus it allows us to recover the articles published when researchers were affiliated to another institution. It also allows us to consider that authors sometimes misreport their institution.

[34] Herself being excluded from the author's lists. Moreover, we only consider the papers with fewer than 50 authors.

[35] This strategy also considers that scientists can publish in different fields

[36] Since reference loops are quite heavy and that this step is likely to run a great number of times for only a few validated articles, we decide to stop it after 30 loops.

12. If some articles are validated in step 11, go to step 2; then stop anyway after two loops;

The similarity between reference lists is based on a score calculated as follows:

$$\alpha_{ij} = \sum_k \frac{1\{i, j \to k\}}{\#\{u\,|u \to k\}},$$

for two papers i and j, of which one is already validated and the other is a candidate paper. The dummy $1\{i, j \to k\}$ takes the value 1 if reference k is cited at the same time by i and j (it is a common reference). The denominator $\#\{u\,|u \to k\}$ is the number of citations that reference k received. It allows us to control the citation frequency of common reference: the more a common reference is cited, the less it should increase the similarity score. We perform the following normalization: $\theta_{ij} = \alpha_{ij}/max_{v=i,j}\{\alpha_{vv}\}$. This normalization is predicated on the maximal similarity that the reference lists of the two papers could reach; that is, the similarity reached if their reference lists were identical, and identical to the one that has the greater self-similarity. The threshold for inclusion is defined as the 98th percentile of all θ_{ij} recorded for the publications of the members in the section.[37]

The collection process is detailed in Table 2.35, which shows the number of retrieved publications and the number of researchers at each stage. The expand stage has been run successively twice, with seven complete loops in the first round and four in the second (each round was followed by 30 loops for the reference list).

Table 2.35 – Number of newly retrieved publications at each step and the number of related researchers

Stage	# documents	# authors
SEED	521,817	29,647
EXPAND		
Round 1		
• Keywords & Authors	585,324	29,309
• References	87,953	29,193
Round 2		
• Keywords & Authors	7,963	29,189
• References	7,929	29,160
Total	1,210,986	29,160
FINAL SAMPLE	1,210,867	29,154

Notes: The column "#documents" gives the number of new validated papers in each step. The column "#authors" gives the number of authors left in the database (authors with more than 500 retrieved papers are removed). At the end of the disambiguation process, a total of 1,210,986 publications is allocated to a sample of 29,160 researchers. Afterwards, the sample is reduced to 29,154 researchers, which equates to 1,210,867 documents, once we correct for homonymy issues.

[37] The 98th percentile was chosen in order to optimize the disambiguation performance.

Benchmarking the disambiguation

This section explains the creation of a benchmark and presents the indicators used to assess the quality of the disambiguation.

We established a list of 353 French researchers who created an ORCID number[38] and were found in our initial list of researchers.[39]

The performance indicators used are precision and recall. Precision measures the ability to clearly identify the correct documents from a set using a common author's identity, whereas recall refers to the ability to retrieve as many relevant publications as possible. These two indicators are scored by:

$$PRECISION = \frac{number\, of\, true\, positives}{number\, of\, true\, positives + number\, of\, false\, positives}$$

$$RECALL = \frac{number\, of\, true\, positives}{number\, of\, true\, positives + number\, of\, false\, negatives}$$

where the true positives stand for the relevant recovered publications, the false positives are the papers retrieved by mistake (they belong to another author) and the false negatives gather the relevant papers that should have been collected, but are missing.

In order to enhance the quality of our disambiguation approach, we decided to set an upper bound of 500 validated publications per author. Hence, at any step of the algorithm, if the number of documents recovered by a researcher exceeds this threshold, we consider that our disambiguation approach has not been relevant enough to treat this homonymy issue, such that the researcher is definitely discarded from the analysis. Once we remove researchers with more 500 articles, we calculate our performance measures from the 291 remaining scientists in the benchmark[40]. We get a recall of 0.90 and a precision of 0.82.

In Figure 2.19, we represent the relationship between the true positives (verified publications) vs. retrieved publications (true positives and false positives) for different measures of outputs. We find that the observations are mostly located on or around the first bisector, which suggests that errors are limited.

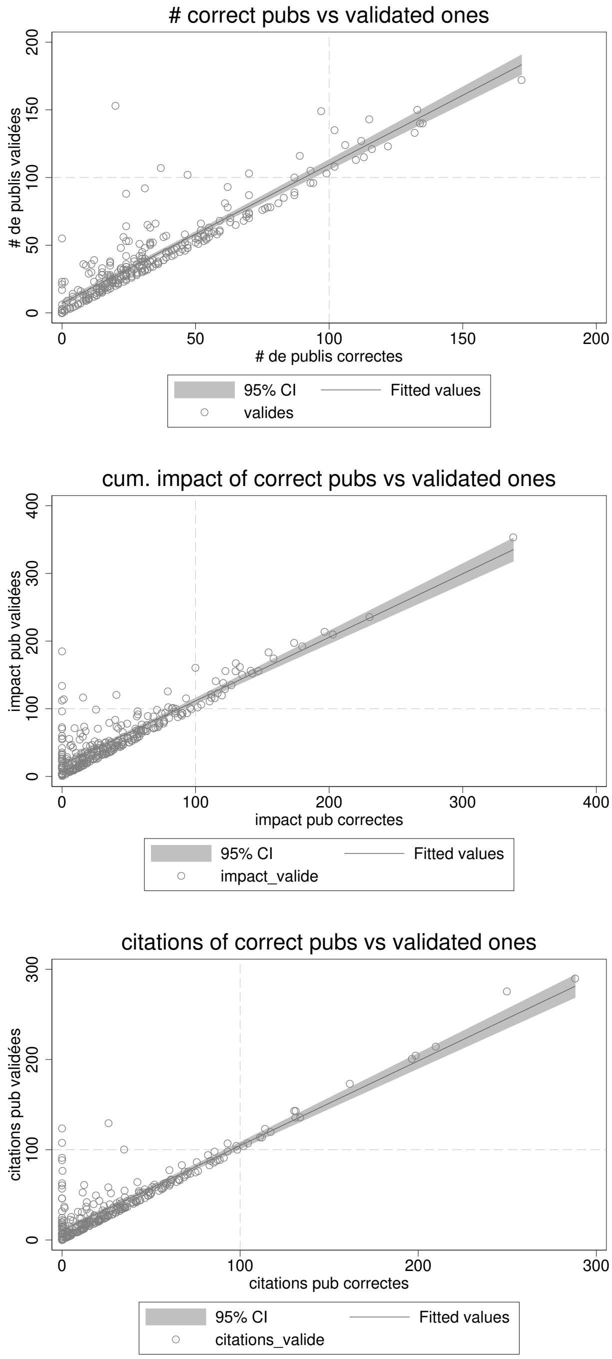
We compare our results with those of Reijnhoudt et al. (2014), who develop a different seed+expand approach to deal with the disambiguation of authors' names. In their paper, the publications collection is based on the similarity between the agents' attributes and some information reported by the articles in the seed stage (affiliation addresses, e-mail addresses), as well as exploiting papers' common features combined with various data sources (WoS, Scopus) in the expand stage. Reijnhoudt et al. (2014) test their method-

[38] An ORCID number lets researchers verify their own publications set on a voluntarily basis.

[39] A manual checking on the similarity of affiliation has been done to ensure there is no homonymy issue.

[40] Among these 291 researchers, no publication was retrieved for 21 of them during the disambiguation process. The mean age of this benchmark is 42.14, and those individuals are affiliated to different research institutes in France. Almost 60% of the sample is affiliated to a university (22% of professors, 37% of associate professors), whereas 40% work as full time researchers.

Figure 2.19 – Comparison of scores on three indicators, comparing correct vs. retrieved measures for the professors and researchers in the benchmark



ology performance on a sample of 1,400 researchers with verified publication records in the period 2001-2010 (“CWTS ’gold standard’”), drawn from a set of 6,753 Dutch full professors. From this subset, they obtained a recall close to 0.96 and a precision in the range of [84.2 - 88.5] for three different versions of the expand stage. At first sight, our study seems to perform to a lesser extent, but this gap may be explained by some differences in design. Firstly, their indicators are calculated on a restricted subset of authors, that is to say those who retrieved at least one publication from the seed stage, whereas we also include those for whom no publication is recovered during the seed stage. As a consequence, their precision is positively biased because false positives are artificially reduced (the recall remains unchanged). Our second remark relates to the selection of individuals to include in the benchmark. In the study of Reijnhoudt et al. (2014), the verified list of relevant publications was systematic, obtained according to the authors’ request or directly from the administration. Our benchmark is different since limited to professors and researchers who created an ORCID profile. Such profile creation is voluntary and unsolicited, which means it is more subject to selection bias. We suspect that the main reason which spurs some researchers to create an ORCID number (and verify their publications) is a large and complicated publication profile which could prove to be difficult to disentangle automatically. Typically, this is the case when authors have been mobile in their career or publish in different scientific domains. Thus they want to clarify the authorship of their publications record, which is facilitated by creating an ORCID profile. As a consequence, the use of our specific benchmark is likely to introduce a negative bias on the recall and precision indicators.

Chapter 3

Implementing Scientific Excellence Via a Site & Cluster Policy. The IdEx Program.[§]

[§]This chapter is based on an ongoing work carried out in collaboration with Nicolas Carayol and Oscar Llopis.

Introduction

As the number of students studying at a university abroad nearly quadrupled from the 90's up to an estimated five million in 2014,[1] the demand for information on university quality at the global level raised significantly in the last decade. This in turn stimulated the emergence/expansion of numerous university or school rankings and league tables. In the light of those new metrics, some national university systems, especially in Continental Europe, appeared as loosing their competitive edge in the global science and higher education competition. Countries such as France or Germany would hardly have one or two universities in the top 100 universities worldwide according to rankings mostly based on scientific performance indicators. Top universities are mostly located in the US, and to a lesser extent in the UK.

In light of these evidence, several governments developed policies targeting research sites and research clusters in order to favor the emergence of "national champions".[2] Such very ambitious science policy is much more complex than project funding policy we have analyzed in the previous chapters of this thesis. They often involve a number of distinct dimensions which ultimately contribute to excellence. In this chapter, we propose a first assessment of the impact of such a policy in France, namely the IdEx Program. The IdEx program aims to favor the emergence of a limited number of world class research universities. Those institutions may preexist to the policy, but most often they are created to comply the goals of the policy, merging distinct institutions located on the same site (a campus or a city). The institutional change implied by the policy goes down to research organization through the creation of research clusters. Whereas top academic institutions have usually been settled in hundreds of years, such an explicit goal of creating top global institutions sounds ambitious. For instance the UK Research Assessment Exercise (RAE), which can be seen as an alternative policy having different aspirations but a similar excellence goal (Moed 2008), involves no specific institutional change. Consistently, there is no forward dimension of the RAE, as funds are coming later out of a backward evaluation. Economics has rediscovered the importance of the quality of social institutions in economic growth (Rodrik 2000, Acemoglu et al. 2012) and at the same time has acknowledged the major role of science in fueling innovation and growth (see Geuna & Rossi (2015) for an overview). But is it really possible to successfully implement such research excellence policy based on institutional change? To what extend precisely does it work? Are all dimensions of scientific excellence equally attainable by such policy? These are some of the main questions addressed in this chapter.

There are several peculiarities of the IdEx initiative that altogether make it particu-

[1] Cf. http://monitor.icef.com/2015/11/the-state-of-international-studentmobility- in-2015/

[2] Such policies were applied for instance in Germany (Möller et al. 2016, Schmoch & Schubert 2009) and in the Nordic countries (Langfeldt et al. 2015).

larly appealing. First, its competitive nature. Unlike block funding, the selection process of the IdEx is based on an open and transparent inter-institutional competition process, which is operated by an independent and international commission. Second, contrary to other research funding schemes that have been previously analyzed (Jacob & Lefgren 2011, Azoulay et al. 2011, Chapters 1 and 2 of this thesis), the funds go to institutions, instead of being assigned to research teams or to individual scientists. Therefore the candidate institutions must prepare their application to the IdEx program across several dimensions (administrative, policy, political, research). Moreover the selected institutions have a high degree of autonomy in deciding the different actions they carry out to materialize the effect of the IdEx policy. Actually increasing the real autonomy of universities is a central goal of the IdEx policy which hinges on the idea that university performance is strongly linked not only to R&D spending but also to good management practices favored by university autonomy (Aghion et al. 2010). This is the reason why each granted university is, at the end of a probation period, given an endowment of up to nearly a billion euros that will guaranty and support its autonomy.

Another important component of the IdEx policy is the formation of research clusters of excellence. Those research clusters most often group together researchers and professors sharing a common field of study, or a common object of investigation from different disciplinary backgrounds, or willing to engage together in an emerging field. Such clusters of excellence were in fact selected through an independent national process prior to the IdEx selection process itself so that, when a research cluster turned out to be part of an IdEx institution, it has been systematically incorporated in that policy. The members of those research clusters not only have their own funds secured, but also are prioritized in most generic policies undertaken by the university. The members of those research clusters thus form the core target of IdEx policy. Not only they are likely i) to receive more research funds, but also ii) to be prioritized by generic university policy and iii) to have their research practices transformed because of the specific research clusters policy. In a sense, they are at the forefront of the policy. They carry the excellence goal of the IdEx policy and therefore constitute a natural "target" to appreciate its effectiveness.
This chapter in effect aims at identifying the impact of IdEx policy on the members of those research clusters.

An issue is that the members of those communities are difficult to identify as membership is "pro facie", and can most often not be inferred using lab data or any available public information. Fortunately, we could obtain this individual membership information for one university, namely the University of Bordeaux and its seven clusters of excellence. Data on the controls (and on the treated in the University of Bordeaux or in other universities in France) are obtained thanks to national lists researchers and professors such as the ones presented in the previous chapters of this thesis. To identify the impact of such core policy, we have identified control individuals who have not been treated by both components of the policy (IdEx and cluster).

As groups of individuals are non-randomly assigned into treatment (IdEx cluster policy), we deal with a clustered observational study characterized by a multilevel structure that we need to take into account. We thus rely on the approach introduced by Zubizarreta & Keele (2017) which allows us to pick similar controls in a two steps matching process (on cluster and individual characteristics). The impact of the treatment is estimated using a difference-in-differences method from the matched sample, to control for the unobserved time-invariant characteristics of the researchers.

We explore simultaneously several dimensions of scientific excellence. The first one is the traditional citation impact that can be approximated by counting the number of highly cited publications (Waltman & Schreiber 2013) or relying on the h-index. The second relates to the originality of research. Clusters of excellence have been selected as they successfully convinced the international committee that they will be able to collectively investigate original lines of research, often at the cross-roads of several disciplines. Does that really work so that the treated eventually pick more novel and original research questions? Do they produce more atypical scientific knowledge? We capture the degree of novelty of research via the atypicality of articles pairwise keyword combinations as suggested by Carayol et al. (2018). Finally, we also intend to appreciate to what extent scientific impact/influence spreads across scientific communities. To explore this latter dimension, we look at the heterogeneity of disciplinary fields that cite the work of scientists (Carley & Porter 2012).

Section I offers an overlook of the IdEx program under investigation. Section II describes the data employed in this study, including descriptions of the main cleaning procedures and matching processes that have been performed prior to the calculation of indicators. It also offers a brief explanation of each indicator as well as the main obtained results. Indicators are categorized into four types: scientific output and impact, scientific excellence, research diffusion and research collaboration. Section III.1 describes the approaches used for the impact analyzes. Section IV presents our results.

I The Excellence Program

At the end of year 2008, the French government decided to launch a large investment plan after the world financial and economic crisis. An ad hoc commission chaired by two previous prime ministers proposed to leverage large funds to make investments sustaining the nation's competitiveness. The whole program was finally launched in 2010, under the label *Plan d'Investissement d'Avenir (PIA)*. In two stages (2009 and 2010), the French governments contributed to this initiative up to a 46 billion euros budget.

We are here interested in only two specific programs of the PIA that took place sequentially and which accumulated budget amounts to 8.15 billion euros.

In the first one, 171 clusters of research excellence were first selected (in two waves, in years 2010 and 2011), by an independent international committee. They received 1.8 billion euros for ten years, that is 10.5 million euro each on average. In the second one, the Initiatives of Excellence (IdEx) program was launched, with a budget of 7.3 billion euros. After a public call for projects, seventeen IdEx proposals where filed. The French Government settled down an independent international commission of experts that were in charge of evaluating each proposal. In July 2011, the international commission selected the first three IdEx for initial funding: Bordeaux, Strasbourg and Paris Sciences Lettres. In February 2012, five additional IdEx were approved: Sorbonne universités, Sorbonne Paris Cité, Saclay, Aix-Marseille and Toulouse. The final economic endowment for this program was 6.35 billion euros. In April 2016, after an initial trial period of funding, the same committee that made the initial selection carried out a second assessment of all IdEx that were selected on the first round. After this second evaluation, the committee decided to finally confirm three IdEx: Bordeaux, Strasbourg and Aix-Marseille. Two IdEx were terminated (Toulouse and Sorbonne Paris Cité). The probation period has been extended for 18 or 24 months for Saclay, Paris Sciences Lettre and Sorbonne Universités.

In the probation period, only the financial returns of the allocated funds could be used. This considerable amount are conceived as starting endowments for those universities, and act, as in US and UK universities, as a guaranty for their autonomy and their capacity to launch original and specific initiatives.

In year 2017, a new stage of the PIA program has been launched which allowed to support additional IdEx and similar initiatives supporting research and teaching institutions. The same independent committee that participated in the first round decided to validate three new IdEx: Côte d'Azur and Grenoble Alpes in 2016 and the University of Lyon in year 2017. Table 3.1 summarizes the information on each IdEx proposal. Other institutions were granted through a variant of the IdEx program called "Science-Innovation-Territoires-Economie" which is focused more on improving connections with society, innovation and teaching quality and less to specifically support the emergence of national champions in research. We mention this 2017 stage for completeness, though this part of the policy is out of the scope of our study. In a sense, we exploit this lag as a number of controls for our treated are employed in these research sites.

In this study, we focus on the University of Bordeaux. As a consequence of its inclusion in the IdEx program, the institution obtained an endowment of 0.7 billion euros. As pre-

Table 3.1 – Initiatives of Excellence: IdEx

University	Start	Amount	Status
PIA 1			
Université de Bordeaux	2011	0.70	Confirmed in 2016
Université Paris Sciences et Lettres	2011	0.75	Probation period extended in 2016 and 2018
Université de Strasbourg	2011	0.75	Confirmed in 2016
Université d'Aix-Marseille	2012	0.75	Confirmed in 2016
Université Paris-Saclay	2012	0.95	Probation period extended in 2016 and 2018
Sorbonne Paris Cité	2012	0.80	Terminated 2016 but reintegrated in probation in 2018
Sorbonne Universités	2012	0.90	Probation period extended in 2016, and confirmed in 2018
Université Fed de Toulouse-Midi-Pyrénées	2012	0.75	Terminated in 2016 *
PIA 2			
Université Côte d'Azur	2017	0.50	-
Communauté Université Grenoble-Alpes	2017	0.80	-
Université de Lyon	2017	0.80	-
Science-Innovation-Territoires-Economie			
Paris-Est	2017	0.28	-
Bourgogne - Franche-Comté	2017	0.33	-
Université de Lorraine	2017	0.33	-
Université Montpellier	2017	0.55	-
Université de Cergy Pontoise	2017	0.28	-
Université de Nantes	2017	0.33	-
Université de Pau et des Pays de l'Adour	2017	0.19	-
Université de Clermont-Auvergne	2017	0.33	-
Université de Lille	2017	0.50	-

Note: Amount in Billions of euros.
* While the projects of Sorbonne Paris Cité and Université Fedérale de Toulouse-Midi-Pyrénées were judged as inadequate and ended in 2016, these group of universities were allowed to submit a new project in 2017. While the Sorbonne Paris Cité project was reintegrated to obtain an extension of its probation period in 2018, the Université Fedérale de Toulouse-Midi-Pyrénées project was not reintegrated and thus has finally terminated.

viously indicated, its funding has been confirmed in 2016 after an initial trial period. As for the other IdEx, the University of Bordeaux could not use this endowment but only its financial returns. This allowed the University of Bordeaux to develop a range of site-level policies as well as to fund research initiatives, in particular the ones of the seven local clusters of excellence which had been selected. Actually, out of the seven filed proposals of Clusters of Excellence involving the laboratories of the University of Bordeaux, five were ultimately granted in 2011. But once the University of Bordeaux received its own IdEx grant, it decided to support the two non-granted research communities at its own expenses. That makes seven clusters of excellence which cover the most renowned scientific communities of the campus, such as Neurobiology, Materials Physics and Chemistry, Ecology, Archeology, etc

II Data and variables

II.1 Research staff

With the help of the University of Bordeaux administration, we collected a research staff list. Up to 5,099 individuals were identified, among which 1,281 tenured researchers (employed by national research institutes working in laboratories jointly run by the university) and 2,803 tenured faculty members. Out of those persons, the managers of the local clusters of excellence identified 991 persons who are part of their cluster.

To identify the impact of the policy, we need a population of controls who were not even partially treated. For this purpose, we first rely upon national lists of professors and researchers associated with laboratories recognized by the Ministry of Higher Education and Research which contain, among other information, the first and last names, status (researcher or teacher - researcher), research and teaching fields, research unit, date of birth and gender. As our lists have been constituted since year 2005 to year 2017, most labs in France have been surveyed this way at least twice. All in all, the data concern 111,615 faculty members, researchers and some engineers and other research staff. When we restrict the list to tenured faculty members and researchers, we have 87,105 persons (29,648 researchers and 57,457 faculty members). This list constitutes actualizes and enriches the list used in the previous chapters of this thesis.

We then collect the scientific publications of the global set of researchers and professors from the Thomson-Reuther ISI Web of Science database (WoS). Based on the surname and initials of the first name of the 87,105 researchers and professors, we initially retrieve more than 7,6 billion documents published until 2017. In order to treat homonymy issues, we apply to these publications a disambiguation process built on a seed+expand approach.[3] Several reasons lead us to remove a part of the individuals during this process; Are deleted those individuals that do not validate eventually any publications, or, on the contrary, those who validated a (too) high number of publications.[4] Finally, using a benchmark composed of 351 researchers for which we know their true publication records, we decide to remove the 6% researchers who initially retrieve the highest number of documents from the WoS in order to optimize the disambiguation quality [5] (i.e. before the disambiguation step). The disambiguation allows us to sort through the documents initially collected from the Wos for our final sample, inasmuch only 17% of the documents are finally considered as authored by our researchers, that is to say 1,373,046 publications.[6]

At the end, our sample consists of 51,003 researchers and professors who published at least one publication from 1999 to 2017, out of which 881 scientists belong to the core research clusters of the University of Bordeaux. They are affiliated to 63 distinct labs. The distribution of the researchers across the 7 clusters is as follow:

- 208 in Amadeus,

[3] The disambiguation process used is the same as the one presented in Chapter 1. See also Chapter 2, Appendix G, p.160 for more details on the process.

[4] We decide to remove researchers who validate more than 500 publications over the period, considering it may signal that the disambiguation process did not suitably work for these individuals' names.

[5] The final removal of these researchers has been decided using two indicators which measure the quality of the disambiguation process (i.e. recall and precision, see Chapter 2, Appendix G, p.160 for a description). Using our benchmark to calculate the indicators, we find that removing the 6% individuals to whom we initially attached the highest number of Wos documents allow us to get a precision of 0.841 and a recall of 0.842.

[6] Our researchers may be the authors of slightly more than 17% of the initial documents collection, considering we only keep articles, revues and letters eventually.

- 406 in Brain,
- 520 in Cote,
- 180 in CPU/Sysnum,
- 128 in Laphia,
- 261 in Lascarbx, and
- 201 in Trail.

As a few researchers are affiliated to two clusters over the period, summing those numbers leads to a number larger than 881.

As we focus on the joint effects of the policies, we need to pick controls in a restricted sample of researchers and professors who are neither affiliated to an IdEx nor to an excellence cluster. Fortunately, in our data set, we do know to which lab and which university each individual is affiliated to. As the Universities treated by the IdEx policy and the labs treated by the excellence clusters policy are public information, we are thus able to define a list of persons who have been treated neither by IdEx policy nor by the research cluster policy. We find $20,063$ potential controls in $1,680$ distinct labs. Table 3.2 synthesizes the numbers for both treated and potential controls.

Table 3.2 – Number of distinct labs and number of professors and researchers among treated individuals and among potential controls.

	Treated Core BU IdEx	Potential Controls Not Idex & Not Cluster
Number of distinct laboratories	63	1,680
Number of researchers and prof.	881	20,063

II.2 The Dimensions of Excellence

We compute a variety of scientific indicators, in order to capture different dimensions of scientific excellence.

II.2.1 Scientific Impact: "Big hits"

We rely on one main indicator to capture the scientific impact of scientists: the number of scientists' academic publications that fall within the top 10% most cited publications in the same research field, in the same year. Percentile-based indicators are based on the idea of looking at the position of a publication within the citation distribution of its field rather than at the actual number of citations of a publication, and have been widely employed to capture research excellence (Bornmann et al. 2015, Waltman & Schreiber 2013).

We also use a second indicator which is more related to the overall reputation of scientists. A scientist has an index h if h of her N_p papers have at least h citations each, and the other $N_p - h$ papers have no more than h citations each. In other words, a scholar with an index of h has published h papers each of which has been cited in other papers at least h times. Thus, the *h-index* reflects both the number of publications and the number of citations per publication. The index is designed to improve upon simpler measures such as the total number of citations or publications. This index has several advantages: (i) it combines productivity with impact, (ii) the necessary data is easy to access in Thompson ISI Web of Science database, (iii) it is not sensitive to extreme values, (iv) it is hard to inflate, (v) automatically samples the most relevant papers concerning citations, etc. The index is related to extremal statistics, which is dominated by exponential density distributions, meaning that high h values are difficult to achieve. The number of cites of the publications of a particular researcher can be represented in decreasing order. The *h-index* of the researcher corresponds to the point where the number of citations crosses the publication order.

The *h-index* has been previously employed for a number of purposes. Several authors have used the *h-index* directly to compare the scientific output of researchers: Hirsch (2005) originally presented a comparison between prominent physicists. González-Alcaide et al. (2008) employed the *h-index* to evaluate scientific research for several authors in different areas in the Biological Sciences and Biomedicine. The *h-index* can also be used to measure the scientific output of complete research groups, institutions and groups of authors (Egghe & Rao 2008, Molinari & Molinari 2008). Van Raan (2006) uses it to compare 147 different chemistry research groups.

Although the use of the *h-index* is well spread, the index suffer from a number of limitations. Some of the critics to the *h-index* that can be found in the literature is that the *h-index* does not take into account several different variables that are often useful to evaluate the production of researchers. For instance, it has been argued that the *h-index* is approximately proportional to career length. That means that the *h-index* is a less appropriate measure of academic achievement for junior academics, as their papers have not yet had the time to accumulate citations. One way to facilitate comparisons between academics with different lengths of academic careers is to divide the *h-index* by the number of years the academic has been active. Thus, an improvement to the *h-index* to compare scientists with different lengths of scientific careers is to divide it by the number of years of research activity (proxied by the number of years since the first paper published by the author).

II.2.2 Scientific novelty

We capture academic novelty based on the frequency of pairwise combinations of author keywords proposed by Carayol et al. (2018).[7] Pairwise keyword combinations capture the different "angles" of a scientific paper. The use of author keywords is convenient as they are freely chosen by the authors and it thus avoids relying on categories defined or redefined by any external actor. Such author keywords have for instance been employed to identify research trends in a range of research fields (Zhang et al. 2016, Li et al. 2009).

To build our indicator, we consider all pairwise keywords combinations of papers published in a given year and research field (identified by the 252 WoS subject categories). We consider keyword combination frequencies within subject categories. Keyword frequencies are computed for the publication year only. Formally, the commonness of the combination of keywords i and j, in subject category c and year t is computed as follows:

$$Com_{ijct} = \frac{N_{ijct}/N_{ct}}{\frac{N_{ict}}{N_{ct}} \times \frac{N_{jct}}{N_{ct}}} = \frac{N_{ijct} \times N_{ct}}{N_{ict} \times N_{jct}}, \tag{3.1}$$

with N_{ct} the number of (non-distinct) keyword combinations in papers published in subject category c and year t. The terms N_{ict}, N_{jct} and N_{ijct} give the number of articles which use respectively keyword i, keyword j, and both keywords i and j. Equation (3.1) manifests itself simply as the share of keyword pairwise combinations that use i and j in the domain c, divided by the expected share of such pairs given the number of times keywords i and j are used in c. Most articles have more than two keywords, which is the minimal possible value. For a given article, let K denote its set of distinct (unordered) pairs of keywords and C its set of associated subject categories. We use the value of the 10th percentile of the distribution of pairwise commonness values for each paper in subject category $c \in C$:

$$com_c = 10thPercentile\left(Com_{ijct} \,|\forall ij \in K\right). \tag{3.2}$$

We then use the inverse logarithmic transformation of commonness to have the novelty of a given paper in a given subject category c:

$$nov_c = -log(com_c) \tag{3.3}$$

Since journals, and thus articles, may be attached to multiple subject categories, we attribute the maximal novelty over all associated subject categories:

$$nov = max_{c \in C}(nov_c). \tag{3.4}$$

[7]Previous and recent contributions have employed the frequencies of pairwise backward references (Wang et al. 2017, Uzzi et al. 2013).

II.2.3 Knowledge diffusion

In order to explore the degree to which the academic knowledge generated by our sample of scientists is spread across academic disciplines, we adopt the idea of the forward diversity index proposed by Carley & Porter (2012). This indicator is forward-oriented, in the sense that it measures the propagation of knowledge based on the citations that a given scientific article has received. For a given publication, we are interested in assessing its diffusion across distinct disciplines. To compute this indicator, we employ the subject categories (SC) of the academic journals where citing papers have been published.

In this study we employ two distinct indicators of knowledge diffusion: Simpson diversity and Shannon. For both index, a higher score is obtained when the forward citations are evenly distributed across scientific fields. Instead, lower scores are assigned when forward citations belong to one (or few) academic fields.

The Simpson diversity index[8] indicates the disparity of different subject categories that are citing a given paper. The index considers the number and distribution of different academic disciplines (Subject Categories) of the research papers that have cited a focal paper. For instance, if a paper has been cited only by other papers from the same research field, the Simpson diversity index would be 0, indicating that the academic impact of this paper is very concentrated into one single discipline. Thus, the application to the Simpson diversity index to forward citations may be useful to reflect how broad (or narrow) is the impact that a given scientist is exerting to the scientific community. The Simpson diversity index is here given by:

$$\text{SimpsonKD} = 1 - \sum_{i=1}^{N} p_i^2,$$

where p_i is the proportion of forward citations emanating from the subject category i, and N is the number of subject categories. The index is calculated based on the stock of publications from period $t_{(min)}$ to t, and citations received in period t to $t+2$.

Shannon Index is a measure of information entropy that focuses on the amount of information necessary for identifying a unit. This index is calculated as the negative sum of the proportion multiplied by the natural log of the proportion for each unit. The measure ranges from 0 to the natural log of N, with N being the total number of units (subject categories). Zero represents complete concentration and the natural log of N indicates a system where all units are exactly equal. Therefore, high values of Shannon index indicate that the citations of a given paper are evenly distributed across a large range of different disciplines. Lower levels of Shannon indicate that forward citations are

[8] The Simpson diversity index is similar to 1-HHI, where HHI is the Herfindahl index

highly concentrated in one or few fields. It is is formally defined as follows:

$$\text{ShannonKD} = \sum_{i=1}^{N} p_i \, ln(p_i),$$

where p_i is the proportion of forward citations corresponding to the i_{th} field, and N is the number of subject categories.

For each scientist, we consider the stock of publications from year $t_(min)$ to t. Then, we consider all citations that this publication stock has received in year t. And then, we compute both indicators of knowledge diffusion (Simpson diversity and Shannon) based on the subject categories of the citing articles. For instance, if all citations that the pool of papers received in year t belong to one single subject category, then both knowledge diffusion scores would be 0 for year t.

II.3 Descriptive Comparative analysis

We use first a very basic approach that allows us to benchmark the trajectories of the treated against the average scientist in their reference group, that is the set of professors and researchers specialized in their field of research and working in France. This allows us to fully control for all the (potentially time-variant) factors affecting the researchers outcomes that would be specific to the field and to France.

We rely upon our full list of 51,122 non-treated professors and researchers to find, for each scientist affected by the core Cluster policy, his/her reference group that we define as all the professors and researchers specialized in the same field of research and working in French academic institutions.

Towards this aim, we have created a unique classification of academic disciplines. This allowed us to classify all our list members in at least one academic discipline. The decision to create our own disciplinary nomenclature is grounded on two reasons. First, each organization employs its own nomenclature of academic disciplines.[9] Since our list is composed of individuals belonging to different organizations, we need to homogenize this nomenclature. Second, the High Council for Evaluation of Research and Higher Education (HCERES)[10] has created its own disciplinary nomenclature, starting in 2012. This nomenclature is based on large disciplines, which pushed us to define our own classification based on large disciplines too. In particular, the nine disciplines of our new classification

[9]For instance, there are specific nomenclatures for CNRS researchers, INRA researchers, INSERM researchers, IRD researchers and faculty members. Moreover, these nomenclatures are not necessarily constant in time (e.g. the nomenclature for CNRS researchers has changed from 2012).

[10]The HCERES is an administrative independent body in charge of evaluation of research and Higher Education in France.

are Biology, Heath, Ecology, Chemistry, Physics, Engineering, Mathematics, Social Sciences and Human Sciences.

For each output indicator that we are going to calculate, say $x_{i,t}$, we calculate the relative difference with a reference group, in the following form:

$$x^{\star}_{i,t} = \left(x_{i,t} - \hat{x}_{i,t}\right) / \hat{x}_{i,t}, \tag{3.5}$$

with $\hat{x}_{i,t}$ the arithmetic mean of the variable considered in the reference group of i in the period t. This indicator thus gives us the difference (in percentage) between the raw individual measurement and the average individual in the reference group.

The reference group is defined as the set of permanent individuals from i's scientific discipline. For the presentation of results, the treatment year is 2011 ($t = 0$).

We report, in a series of figures (Figures 3.1–3.5), the evolution of the mean treated relative to its reference group (treated or not) for the number of publications and each of the above-mentioned variables of interest which reflect different dimensions of scientific excellence.

Interestingly, it appears that the publication counts respond very positively after treatment, the h-index has a positive but lower trend after treatment, and the number of big hits has a negative trend post treatment. Novelty remains stable but is negative over the whole period (articles of treated researchers are less novel than the average in their field). The diversity of citing sources is positive and raises again after a declining period centered around the treatment year.

At this stage we need to be very cautious before trying to read into those results. It is very important to recall that a certain number of persons in the peer groups have also been treated by the similar policy, in other universities. Moreover, it is impossible at this stage, to ascertain that the observed evolutions post-treatment can be explained as direct consequences of the participation in the program under evaluation, or whether it could be explained as a consequence of other factors. For this purpose, we conduct an impact analysis in the next subsection, which allows us to built a control group that is not only similar to the treated in terms of the research discipline.

Figure 3.1 – Relative number of publications of the Core Bordeaux IdEx community (research clusters). In the right graph, article counts are adjusted for coauthorship.

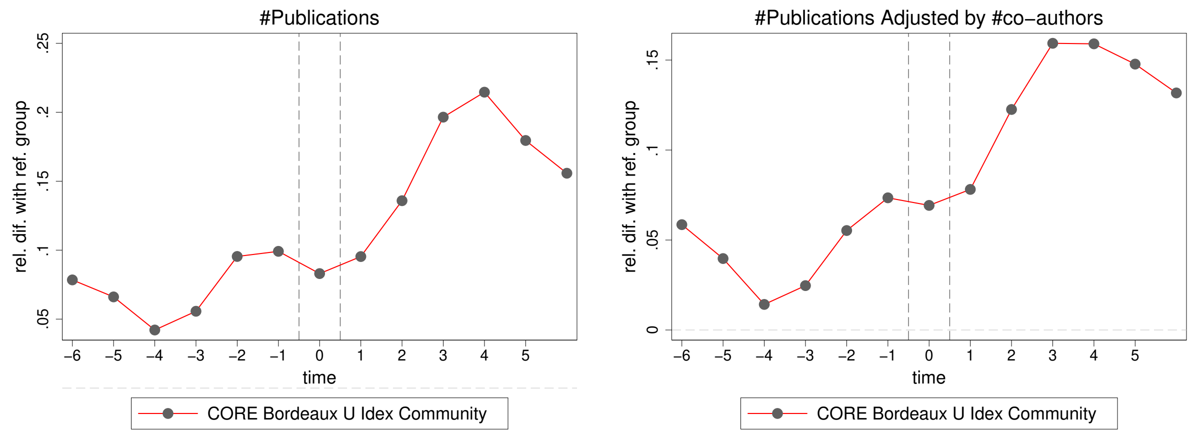


Figure 3.2 – Relative number of top cited articles (top 10%) in each field of the Core Bordeaux IdEx community (research clusters). In the right graph, the proportion of such top papers is considered instead of the mean.

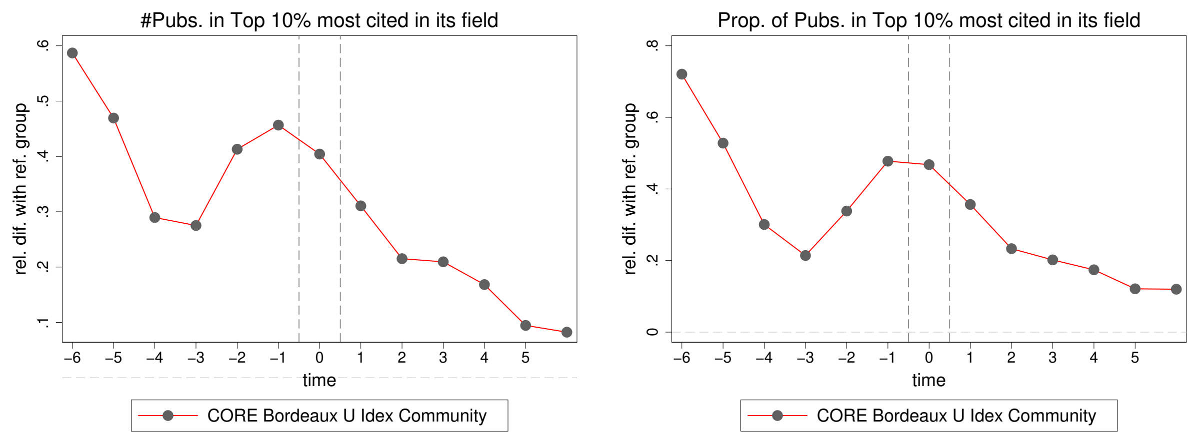


Figure 3.3 – Relative h-index of the Core Bordeaux IdEx community (research clusters). In the right graph, the h-index is adjusted for age.

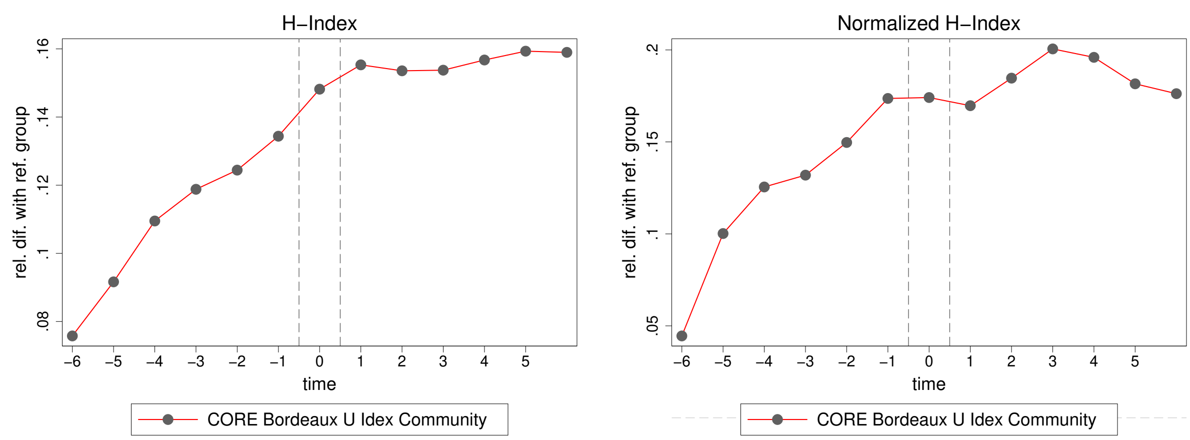


Figure 3.4 – Average novelty of articles published in the Core Bordeaux IdEx community (research clusters). In the right graph, the maximum novelty (instead of the mean) is considered for each author.

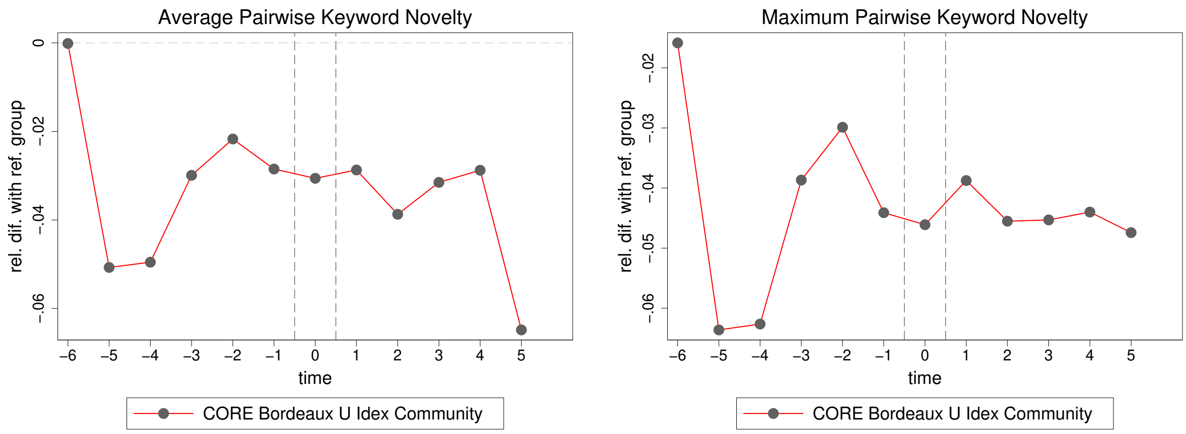
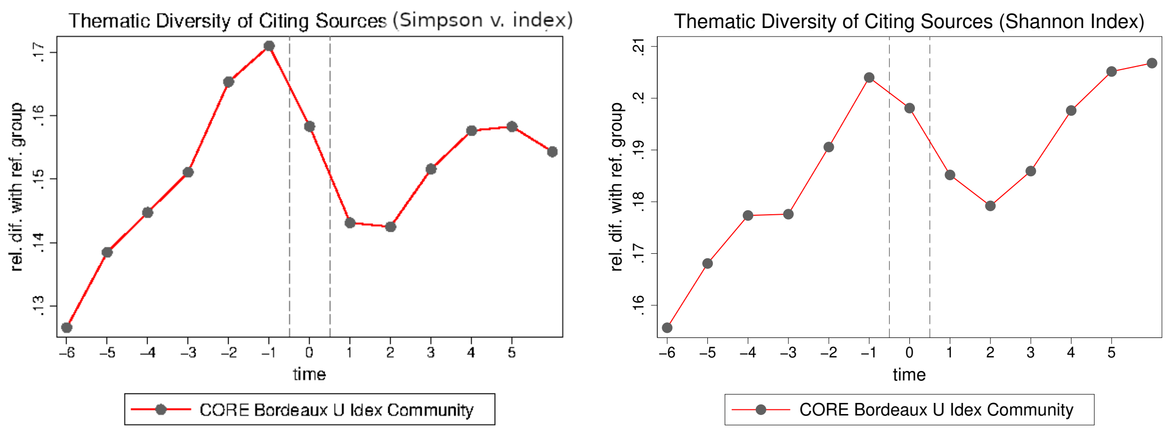


Figure 3.5 – Relative diversity of the citing sources of the publications of the Core Bordeaux IdEx community (research clusters). The left graph makes use of the Simpson diversity index while the right graph use the Shannon index.



We report similar investigations in Figures 3.6–3.10 for the evolution of the factional polynomial estimates of the same variables. In Figures 3.11–3.15, we report the histograms calculated on years before treatment (2006-2009) and year after treatment (2013-2016).

Figure 3.6 – Fractional Polynomial estimates of the relative number of publications of the Core Bordeaux IdEx community (research clusters). In the right graph, article counts are adjusted for coauthorship.

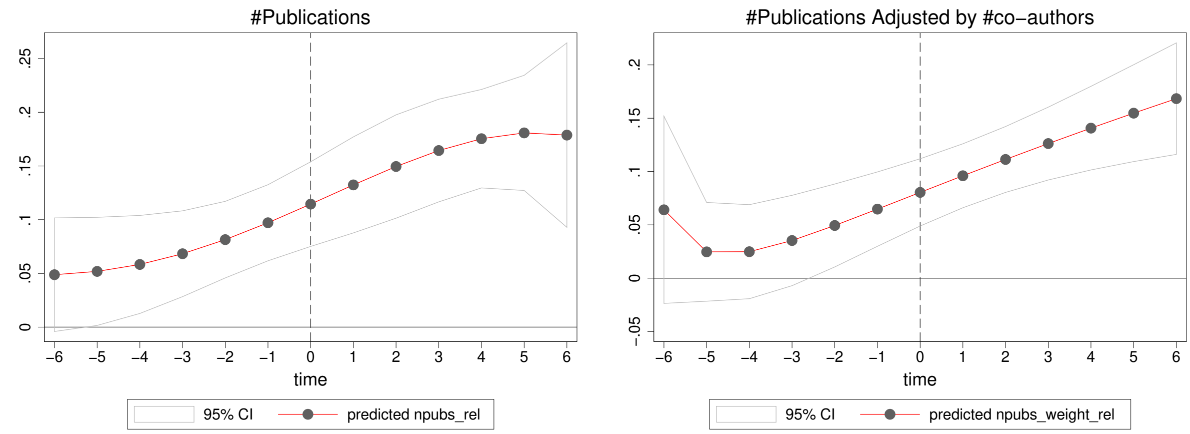


Figure 3.7 – Fractional Polynomial estimates of the relative number of top cited articles (top 10%) in each field of the Core Bordeaux IdEx community (research clusters). In the right graph, the proportion of such top papers is considered instead of the mean.

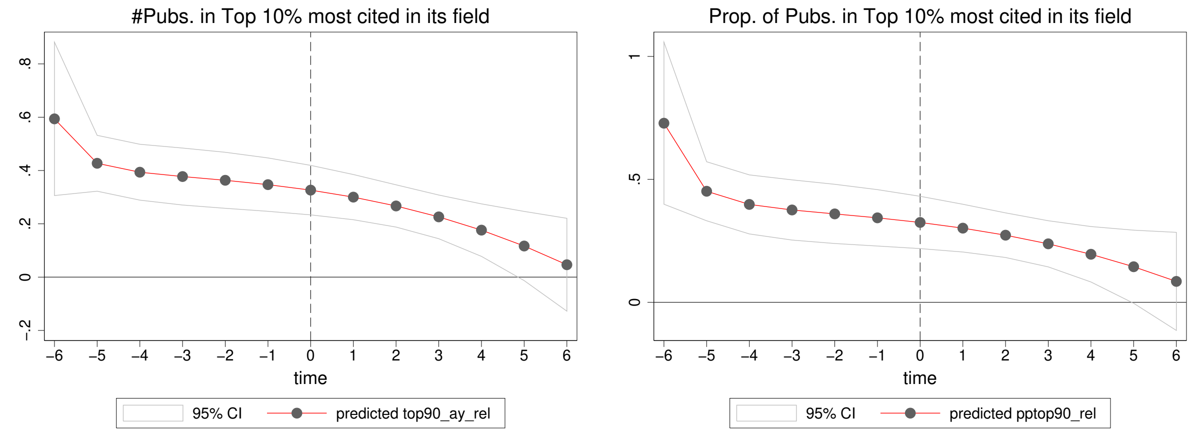


Figure 3.8 – Fractional Polynomial estimates of the relative h-index of the Core Bordeaux IdEx community (research clusters). In the right graph, the h-index is adjusted for age.

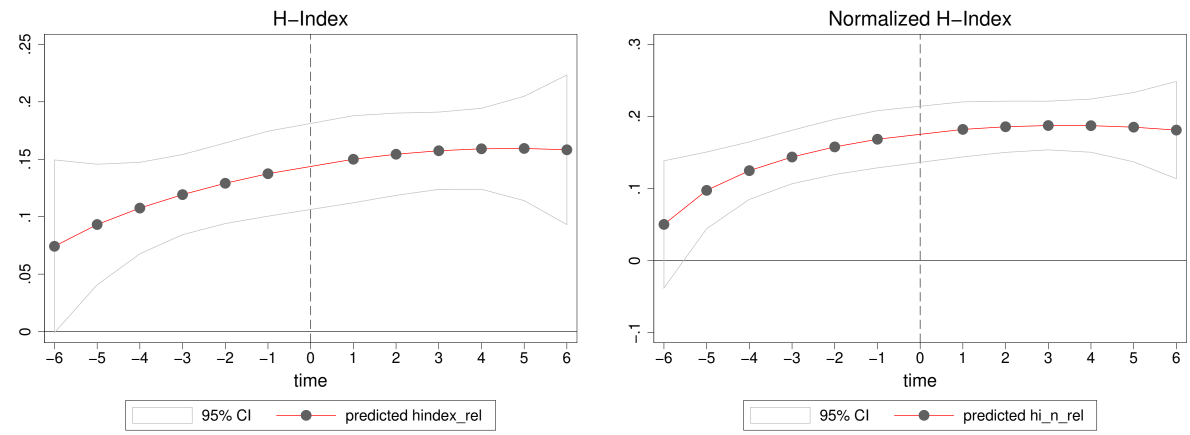


Figure 3.9 – Fractional Polynomial estimates of the average novelty of articles published in the Core Bordeaux IdEx community (research clusters). In the right graph, the maximum novelty (instead of the mean) is considered for each author.

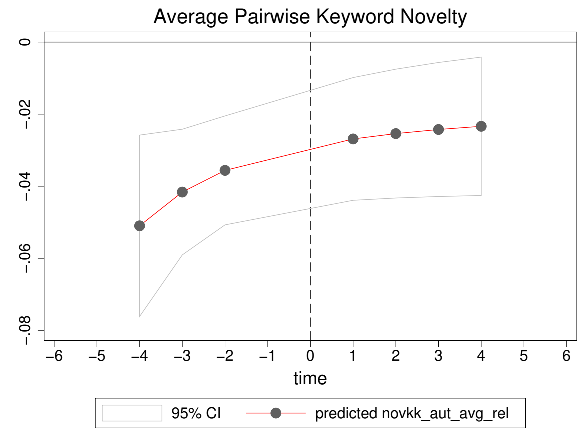


Figure 3.10 – Fractional Polynomial estimates of the relative diversity of the citing sources of the publications of the Core Bordeaux IdEx community (research clusters). The left graph makes use of the Simpson diversity index while the right graph use the Shannon index.

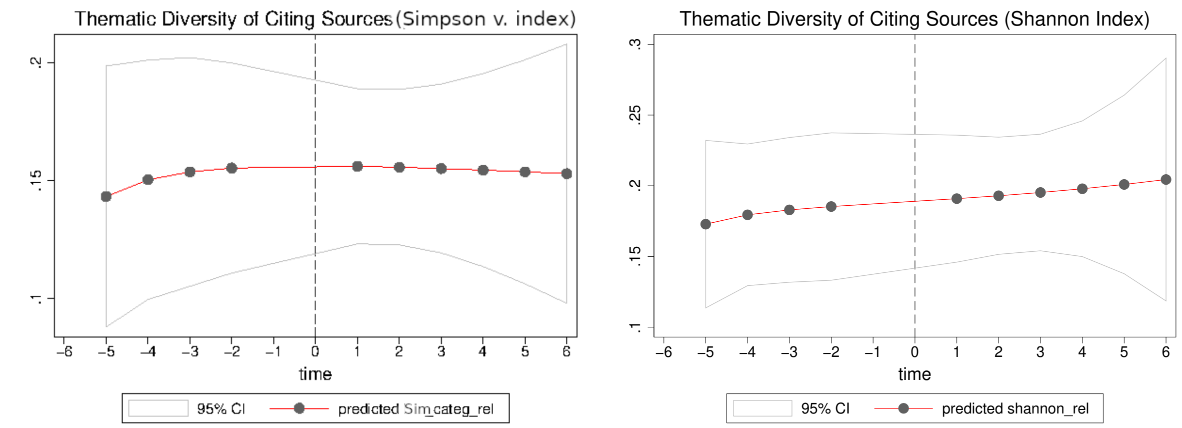


Figure 3.11 – Histograms of the relative number of publications of the Core Bordeaux IdEx community (research clusters), before and after the treatment. In the right graph, article counts are adjusted for coauthorship.

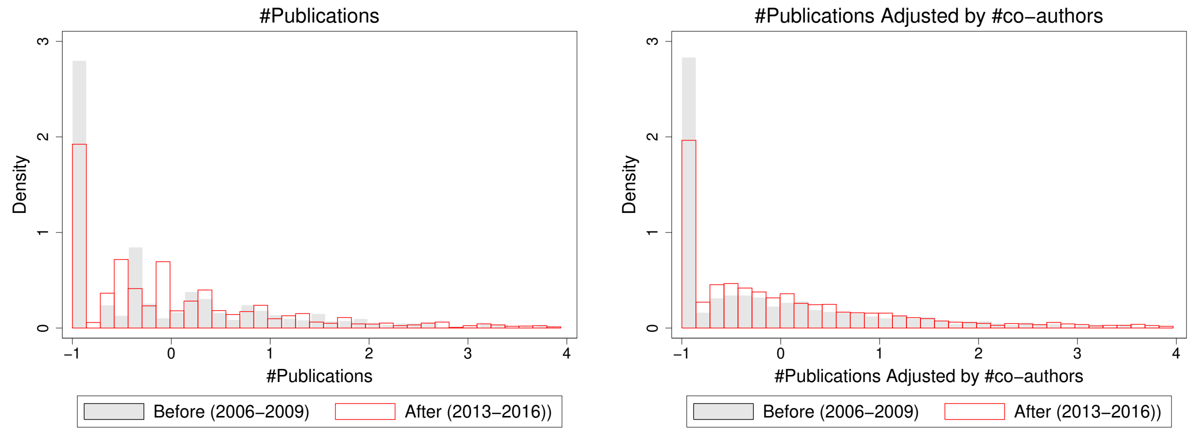


Figure 3.12 – Histograms of the relative number of top cited articles (top 10%) in each field of the Core Bordeaux IdEx community (research clusters). In the right graph, the proportion of such top papers is considered instead of the mean.

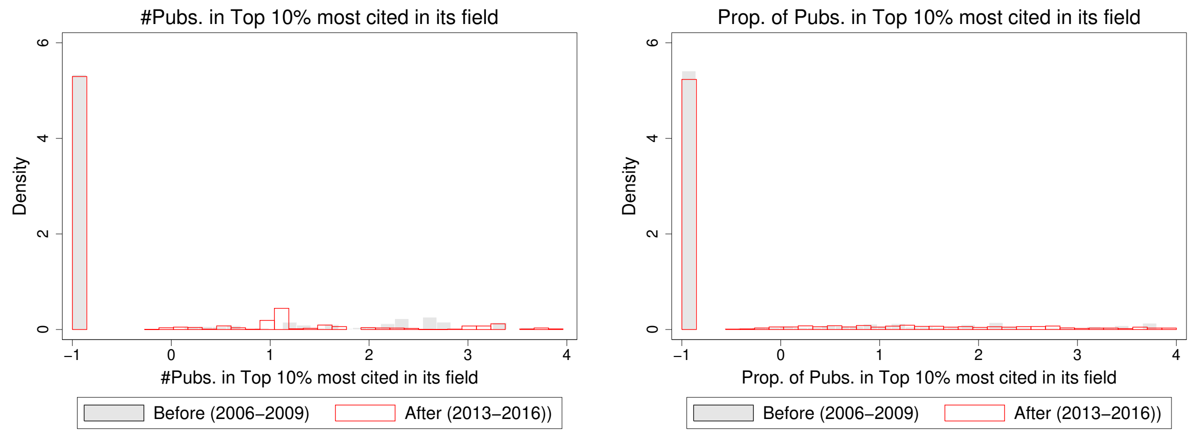


Figure 3.13 – Histograms of the relative h-index of the Core Bordeaux IdEx community (research clusters). In the right graph, the h-index is adjusted for age.

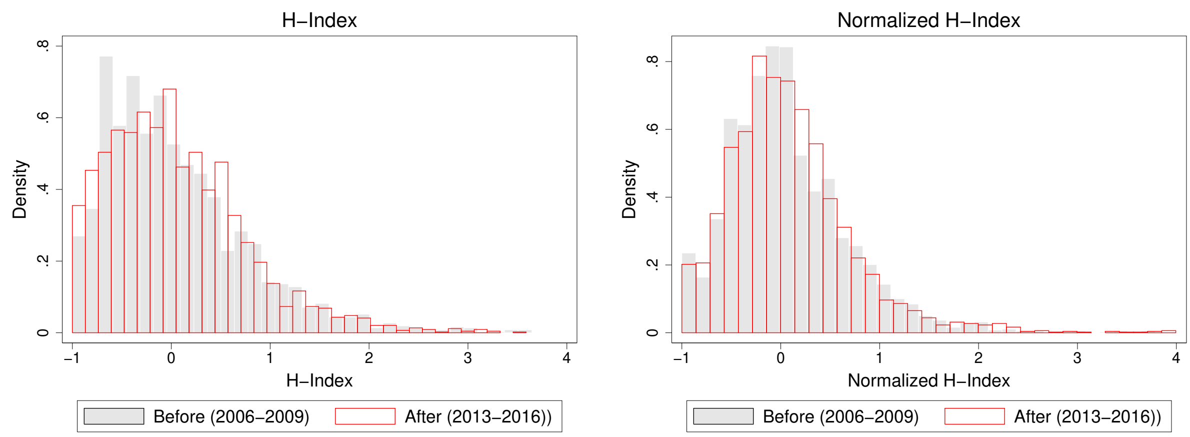


Figure 3.14 – Histograms of the of the average novelty of articles published in the Core Bordeaux IdEx community (research clusters). In the right graph, the maximum novelty (instead of the mean) is considered for each author.

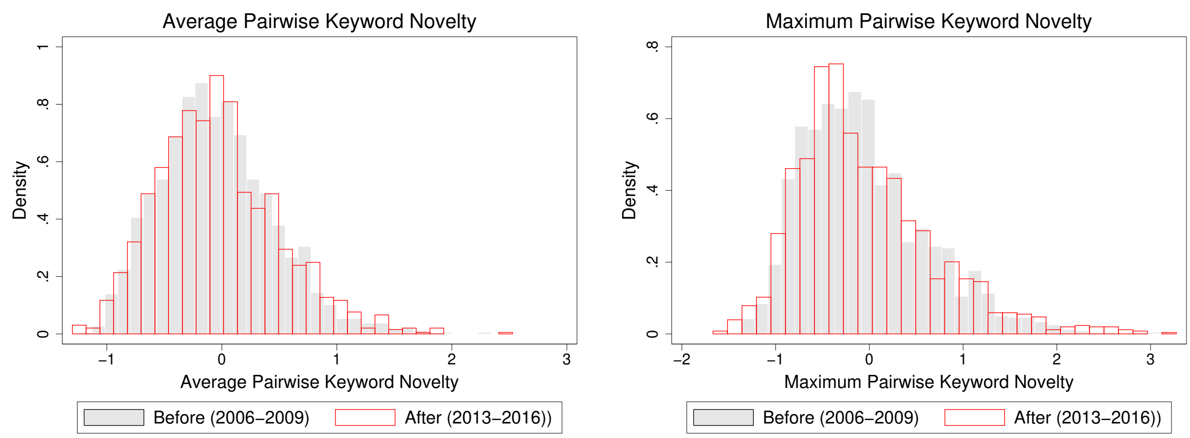
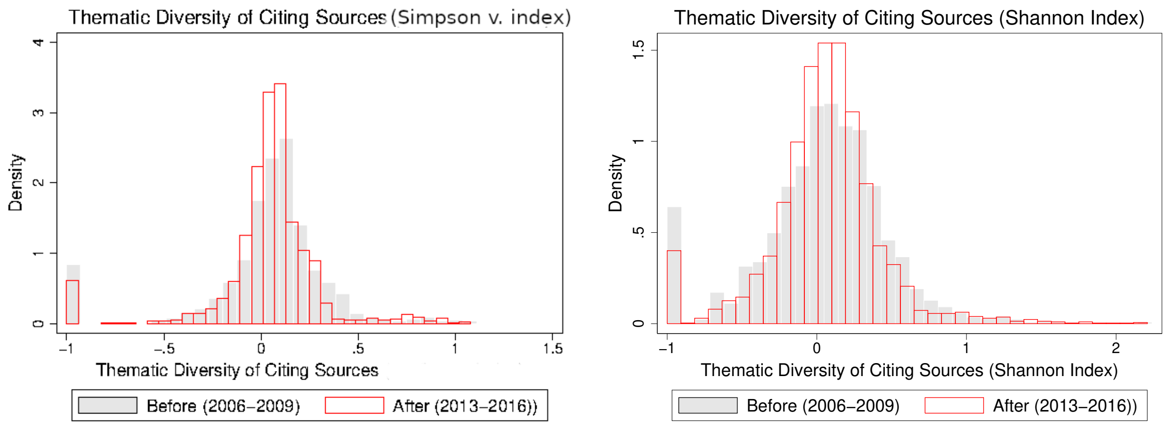


Figure 3.15 – Histograms of the relative diversity of the citing sources of the publications of the Core Bordeaux IdEx community (research clusters). The left graph makes use of the Simpson diversity index while the right graph use the Shannon index.



III Identification Methodology

In most non-experimental observational studies, the non-random assignment of units to treatment introduces a bias in impact assessment if not specifically considered. This bias is due to the difference between treated and control units in terms of their baseline covariates that act as confounding variables. Relying on matched samples may bring the results closer to the ones which would be obtained had the treatment been randomly assigned conditional only on observed differences between treated and controls.
Different matching methods have been used in the literature. Some require exact matching between units on a small numbers of variables, whereas others require a close matching based on the estimation of propensity scores on a large set of covariates (Rosenbaum & Rubin 1983). Propensity scores are defined as a the estimated probability for a unit to receive the treatment conditional on its covariates. Other approaches rely a close match on a large set of covariates so as to minimize the distance between treated and control units, without directly using propensity scores (Cochran & Rubin 1973, Zubizarreta et al. 2014).

In this study, we have to also pay special attention to the clustered nature of the assignment that may confound treatment. We are in particular worried in the fact that the scientific excellence of the research clusters community explains treatment and is correlated to individual and lab quality which in turn affect future outcomes. Let's recall that research clusters communities are essentially all or parts of research laboratories located in the same geographical area which formed a coalition to set up a proposal. As we aim to appreciate the impact of increased funding in a context of reshaping institutions and research management, in essence we would like to compare the observed outcomes of the treated with the ones of similar individuals acting in a research environment similar to the treated in the absence of treatment. The counterfactual environment is of course not similar to the clusters of excellence as such groups likely do not exist in the absence of funding. Scientists then likely remain in their labs. As peer effects likely exist in research labs (as in most social environments) and as peers' quality in the lab likely correlates with the odds of treatment, our study design needs to take this into account.

Situations concerned with a multilevel structure of the data are ubiquitous. For instance, in education, students are nested in classrooms or schools. In health sciences, patients are clustered by doctor or by hospital. In many observational studies, treatments are assigned at the cluster level rather than at the individual level. As treatment is often not-random, not only the characteristics of the persons matter in these clustered observational studies but also the characteristics which describe the common environment, measured at the cluster level, since both information may confound the causal effect if not controlled for. Different approaches have been used in the literature to evaluate the effect of a treatment in multilevel settings. Some of them rest upon the estimation of the so called "propensity score" and take into account the hierarchical structure of the data (Arpino & Mealli 2011, Li et al. 2013).

Kim & Seltzer (2007) suggest to select the control units within the same group, in their case the school level, to ensure that treated and control units are affected by the exact same environment. While matching within the same group has the advantage to control unobserved group-level variables (in our case, this would amount to match within the same lab with individuals who are not assigned to treatment), we do not follow this approach for several reasons. First, if all the members of a laboratory are not assigned to an Excellence Cluster all of them are treated by the Excellence University policy. Therefore, it can not be used to assess the effect of both policies simultaneously as we need to select controls treated by none of the policies. Secondly, constraining the matching to occur in the same laboratory harshly limits the number of potential controls to build the control group.

We apply the methodology introduced by Zubizarreta & Keele (2017) which allows us to match on both individuals and groups of individuals (labs). Our objective is to find controls for both individuals and their lab environment. For each laboratory participating in a cluster, we seek to find a similar untreated laboratory, and the pool of these control laboratories would represent the counterfactual outcome of our treated research clusters. These control labs have no link to any research cluster of excellence nor to an IdEx. Both the individual and laboratory levels are considered in the matching method, and comparison between treated and controls is made at the two levels simultaneously in order to optimize the matched sample. We count the individual scientists and laboratories that may act as a control in Table 3.2.

III.1 Matching method

We are interesting in the estimation of the average treatment effect on the treated (ATT). We can define it from the potential outcome framework developed by Rubin (1974), such as:

$$ATT = E\left(Y(1) - Y(0)|\,T = 1\right), \tag{3.6}$$

where $Y(1)$ denote the outcome variable observed whether the unit is assigned to treatment, and $Y(0)$ represents the outcome when the unit is not treated. Given that the unit receives indeed the treatment ($T = 1$) and that both potential outcomes are not observed simultaneously for a particular unit, we rely on the control group to recreate the counterfactual outcome of the treated, $Y(0)$.

One central hypothesis that need to be valid, unless not testable, is the Unconfoundedness Assumption which implies that all individual-level and lab-level confounding covariates are observed and are balanced between the treated and control groups. The hypothesis of Weak Unconfoundedness,[11] required for the estimation of the ATT, can be expressed as follow:

$$Y(0) \perp T|X. \tag{3.7}$$

In words, it means that the potential outcome associated to the situation of not being treated by the policy $Y(0)$ is required to be independent of the treatment status T conditional on the set of confounders X. In addition, the probability to receive the treatment, $P(T)$, has to be identifiable for the treated, with $P(T) < 1$.

The objective of the two-step matching method used for the selection of controls is not to exactly replicate a randomized controlled trial (CRT), in which the treatment is randomly assigned to a group of units. Indeed, in this setting, group-level covariates are balanced, but unit-level covariates are not necessarily similar (Hansen & Bowers 2009). The goal is rather to mimic a paired group randomized controlled trial, in which groups are matched on baseline covariates before the clustered treatment assignment. Our approach involves cardinality matching process, which is applied to matches at different levels (researcher and laboratory levels in our case). The optimal matching is implemented by one to one matching without replacement between treated and control units, that is to say that each treated researchers affiliated to a particular laboratory is matched to one relatively similar not treated researcher from a paired laboratory. The matching

[11] This assumption is less restrictive than the Unconfoundedness assumption stated for the estimation of the Average treatment effect (ATE), and expressed as $Y(0), Y(1) \perp T|X$. In this case, the potential outcomes $Y(1)$ and $Y(0)$ need to be independent of the true treatment status T conditional on the confounding variables X. A second hypothesis rely on the possibility for each observation to be treated, with $0 < P(X) < 1$ the common support assumption.

is only performed on a set of individual covariates, but the method permits inclusion of additional constraints on laboratory level covariates, that would be balanced using a fine balance method (Rosenbaum 1989, Rosenbaum et al. 2007). This fine balance process implies the marginal distribution of the variables to be balanced between groups. These group-level variables are required to be nominal and are used alone, as well as interacted among themselves in the balancing constraints. The underlying algorithm has the advantage of automatically implementing the covariates balancing requirements, when possible, between treated and control pairs for each level of matching (Zubizarreta & Keele 2017, Pimentel et al. 2018). The importance of covariates balance, as a way to follow the Unconfoundedness assumption,[12] has been pointed by several authors (Rosenbaum & Rubin 1985, Dehejia & Wahba 1999, Rosenbaum et al. 2007). The disadvantages of the matching method involve the systematic removal of treated units when potential controls researchers or labs are not enough close to the treated. This multilevel cardinality matching is intended to find the optimal set of combinations in terms of treated/control paired units and ensure to exactly or approximately balance the specified covariates between groups.

The cardinality matching method used works as follow. First, let L_T with $T = 1, ..., N_{lt}$ denote the laboratories to which, researchers involved in research clusters, are affiliated, and let L_C with $C = 1, ..., N_{lc}$ represent the laboratories not implied in a cluster. Then i_{lt} represents the treated individual i affiliated to the treated laboratory L_T and i_{lc} denotes a potential control individual treated neither by the IdEx policies nor by the excellence cluster one. For all the possible combinations of treated and control labs, (L_T, L_C), a distance matrix is built for each pair of individuals i_{lt} and i_{lc} within the pairwise laboratory match. The matrix is built on a defined set of individual covariates, and a lower score is assigned to pairs of units that do not satisfy covariates balance. Low number of good matches within pairs of laboratories are also penalized. Then, using scores derived from these matrix distance, an optimal matching is applied to select the largest number of matched pairs of laboratories among the possible combinations (L_T, L_C). Each laboratory can be observed at most once in the matched sample. At this step, fine balance constraints can be added to specify that some laboratory-level covariates are expected to be balanced in the matched sample. Once laboratories are paired, the largest number of treated researchers i_{lt} affiliated to the laboratory L_T is also matched to close control researchers i_{lc} affiliated to the laboratory L_C, which has been matched to L_T, in a one-to-one matching. The goal of this process is to maximize the size of the balanced researchers pairs within matched pairs of balanced laboratories.

Once a control group composed of individuals who do not differ from our treated is found, we estimate the average treatment effect on the treated by difference-in-differences

[12] This assumption, defined above, implies that the matching process allows to limit differences in observable pre-treatment covariates between treated and control units.

(DID). This method allow us to control for all unobserved covariates which are constant in time, by differentiation between the periods observed before and after the treatment assignment. The ATT is estimated by:

$$\hat{\delta} = \frac{1}{|N_T|} \sum_{i \in N_T} \omega_i \left(Y_{i,1} - Y_{i,0}\right) - \frac{1}{|N_C|} \sum_{j \in N_C} \omega_j \left(Y_{j,1} - Y_{j,0}\right), \tag{3.8}$$

where N_T denotes the set of treated individuals and N_C the set of controls. $Y_{i,t}$ is the outcome variable observed in period t, with $t = 1$ in the period after the treatment assignment, and $t = 0$ in the period before treatment. In a one to one matching without replacement setting, the weight given to each treated unit ω_i and the weight assigned to each selected control ω_j both equal one.

This method is based on the parallel path hypothesis, which states that treated units would have experienced a similar (parallel) evolution in their outcome variables whether they would not have been treated.

In the next subsection, we present the variables used to perform the multilevel matching.

III.2 Matching variables

All the variables that affect both the clustered assignment of treatment and the outcome variables have to be considered for the matching, so that their distribution between treated and control groups would end out being balanced, after the match. The optimal matching method that we use is based on a set of individual covariates that are used to match pairs of treated and control laboratories and then pairs of researchers. These covariates are presented in Table 3.3. As age may affect both the odds of being granted and scientific outcomes, it will be included among matching variables. The same applies to the status of the researchers. We define a dummy variable which differentiate between full time researchers (or associate researchers) and professors (or associate professors). Several scientific production measures are also included: The number of publications in the last three years[13] to assess recent research intensity; the number of citations over the same period in order to account for the direct scientific impact of recent research; and the age-normalized h-index to account for their scientific impact. In the perspective to compare

[13]Scientific production measures are calculated on the three-year window from 2007 to 2009. We voluntary omit year 2010 since researchers who participate in Cluster projects diverted time away from their research activity to prepare the proposal (whereas 2011 refers to the year of the Cluster selection and is naturally ignored).

Table 3.3 – Description of the individual-level covariates used in the matching process and the laboratory-level covariates

Matching covariates	Description
Age	Age in 2010
Status	Dummy variable of the status of the scientist ('1' if assistant professor or full professor, '2' if assistant researcher or research director)
Articles	Mean number of articles published between 2007 and 2009 (fractional count)
Citations	Mean number of citations received by articles published between 2007 and 2009 (fractional count)
Norm H_index	H-index (measured in 2009, age normalized)
Scientific field	Discipline of the researcher ('1' Humanities '2' Social Sciences '3' Mathematics '4' Physics '5' Universe Sciences '6' Chemistry '7' Engineering ('8' Information and Communication Sciences and Technologies '9' Agronomics and Ecology '10' Life Sciences and Environment '11' Microbiology '12' Neurology '13' Physiology and Endocrinology '14' Public health)

Laboratory-level covariates	Description
Article lab	Categorical variable indicating in which quartile is situated the laboratory in terms of the total number of articles published between 2007 and 2009 (fractional count) compared to the total distribution of laboratory of the same main discipline
Max citations lab	Categorical variable indicating in which quartile is situated the laboratory in terms of the maximum number of citations received for an article published between 2007 and 2009 (fractional count) compared to the total distribution of laboratory of the same main discipline
Lab size	Categorical variable indicating in which quartile is situated the laboratory in terms of its size (calculated as the sum of permanent researchers and professors)
Main discipline	Main discipline of the lab ('1' Biology '2' Health '3' Ecology '4' Chemistry '5' Physics '6' Universe Sciences '7' Engineering '8' Mathematics '9' Human and Social Sciences)

researchers who carry out related scientific activity, we also include the scientific field in the set of individual covariates. For each of these covariates, we verify that there is an overlap between their distribution in the treated and control groups. Otherwise, it would violate the common support assumption, that is necessary to match treated to similar controls.

As the characteristics of the laboratory may also influence both the probability of assignment to treatment and outcome variables, we intend to balance relevant laboratory-level covariates. They should account for the scientific environment and the dynamism in the laboratory. We thus construct categorical variables based on the average number of articles published in the last three years and the maximum number of citations obtain by a member over the same period to account for the presence of renowned researchers. We also define two other lab variables. First we create a categorical variable related to the size of the laboratory (in terms of the number of permanent researchers and professors), and we deduce the main scientific discipline of the laboratory from the disciplines the most frequently associated to scientists in the lab. We do not include additional fine balance constraints on laboratory-level covariates since, as showed in the next subsection, our main covariates are balanced in means in the matched sample.

In the matched sample, we still find the 60 different laboratories where treated researchers are affiliated to. However, a sizable number of treated researchers have been removed (-60%), certainly due to a lack of similar control researchers in the matched control laboratories. The matched sample consists of 350 treated researchers along with their 350 controls, similar in pairs in terms of the covariates used for the matching.

III.3 Balance diagnostic

To measure how similar are matched labs and units after matching, a balancing test is applied to the variables describing the characteristics of the individuals and those of the laboratories and is measured by absolute standardized differences of means between the treated and control groups:

$$\text{std.bias} = \frac{\bar{x}_{T=1} - \bar{x}_{T=0}}{\sqrt{\left(s^2_{T=1} + s^2_{T=0}\right)/2}}, \tag{3.9}$$

where the numerator denotes the weighted mean difference between the treated ($T = 1$) and the selected control units ($T = 0$) for the covariate x and is divided by its pooled

Table 3.4 – Standardized difference of means between treated and control groups, before and after the matching

Individual-level covariates	Stand.bias Before	Stand.bias After
Age	-0.273	-0.034
Status	0.407	0.07
Norm H-index	0.525	0.06
Articles	0.275	0.043
Citations	0.323	0.038
Scientific field:		
Humanities	-0.251	0.01
Social Sciences	-0.364	0.00
Mathematics	0.344	0.01
Physics	-0.008	-0.014
Universe Sciences	-0.015	0.00
Chemistry	0.175	0.00
Engineering	-0.212	0.00
Information and Com. Sc. & Tech.	-0.053	-0.01
Agronomics & Ecology	-0.045	0.00
Life Sciences & Environment	0.169	0.00
Microbiology	-0.133	0.00
Neurology	0.353	0.00
Physiology & Endocrinology	-0.088	0.00
Public health	-0.126	0.00
Laboratory-level	**Stand.bias Before**	**Stand.bias After**
Articles lab	-0.284	0.068
Max citations lab	-0.293	0.065
Lab size	-0.067	-0.301
Main discipline of the lab:		
Biology	0.30	0.374
Health	-0.15	-0.401
Ecology	0.186	0.126
Chemistry	0.383	0.155
Physics	0.194	0.00
Universe Sciences	-0.067	-0.103
Engineering	0.006	0.146
Mathematics	-0.096	-0.24
Human and Social sciences	-0.598	-0.084

Notes: The unmatched sample is composed of 18,847 researchers and professors nested in 1,686 different laboratories (817 treated in 60 labs and 18,030 controls in 1,626 labs). The matched sample consists in 650 scientists (325 treated and 325 controls) affiliated to 120 different laboratories (60 labs for treated units paired to 60 other labs with control units).

standard deviation in the unmatched sample. Given that we use a one to one matching without replacement, all weights equal one.

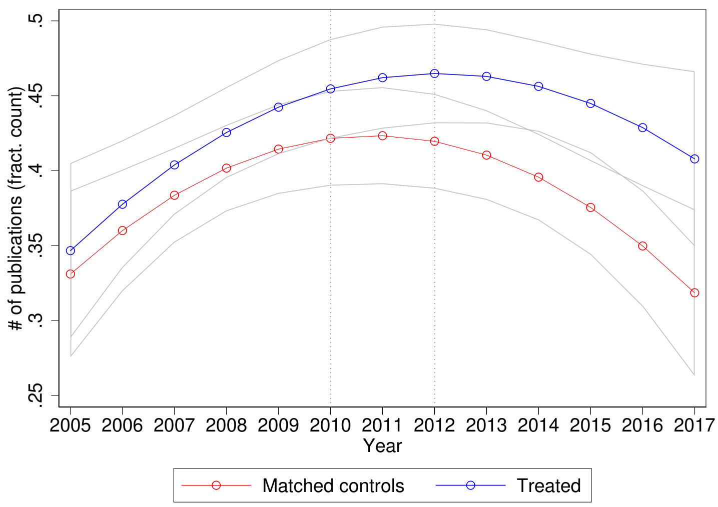
We measure standardized differences of means for each individual or laboratory-level covariate, which may confound the causal effect, both before and after matching (Table 3.4). It is usually stated that the size of the difference have to be lower than 20%, and is appropriate when lower than 10%, which is the expected difference size in randomized experiments (Rosenbaum 2010). We observe that the matching allows us to substantially reduce imbalance between the treated and control groups. The means of all individual covariates used for the matching are strongly balance, since they differ by less than 0.1, that is less than one-tenth of a standard deviation. Although we did not specify any additional constraints on the laboratory-level covariates, the categorical variables created from the mean number of articles published in the laboratory and the maximum number of citations, present a low discrepancy in the remaining bias after matching (less than 7% of a standard deviation). However, the search for bias reduction for our priority covariates during the matching process makes the balance worse for other less important covariates. In fact, the matching increased the difference between groups in terms of the size of the laboratory, given that the bias increased from 7% before matching to 30% of a standard deviation in the matched sample. We also see that a substantial bias remains for several laboratory disciplines after matching (bias may be more than 20%).

IV Preliminary Results

IV.1 Impact analysis

We estimate the ATT on our outcome measures using a difference-in-differences method in order to adjust for unobserved time-invariant confounders. As we can see in Figures 3.16 and 3.17, this approach is necessary since the average outcomes of the treated researchers are often above those of their controls in the period prior to the treatment assignment. A simple difference between groups on the mean outcomes observed in the period after the treatment assignment would have over-estimated the treatment effect. This means that the matching procedure allows to select close but not similar controls for the treated, as showed by the gap between controls and treated before treatment (before year 2011). Therefore, pre-treatment differences seem not to have been totally removed with the matching. However, if the parallel path hypothesis is valid, we may expect that these differences are removed when differentiating with the difference-in-differences method. Considering we observe a six-years period following the assignment to treatment,

Figure 3.16 – Fractional polynomial estimates of the number of publications (Adjusted by # coauthors) of the Core Bordeaux IdEx community (research clusters, in blue) and for their selected controls (in red).

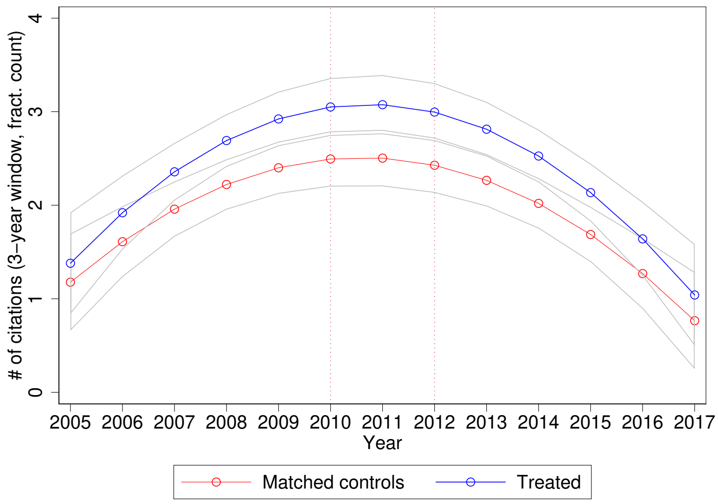


Notes: The observed decrease in the yearly number of publications in the second period can be related to the delayed referencing of recent articles in the WoS database.

we define two outcome measures on this period; a first variable is built to assess the mean effect on the three first years after the setting-up of the research clusters, and a second one for the three next years.

Results are presented in Table 3.5. We first find that the policy implementation has a positive effect on the number of articles published, with a larger coefficient reported for the distant period ($[t+4\ ;\ t+6]$ in column two) but that this effect does not significantly differ from zero. We then turn to the investigation of the impact on our three dimensions of scientific excellence. First we find that the policy does not significantly affect our measures of scientific impact. The coefficient associated to the number of top 10% cited papers is positive but does not significantly differ from zero. We also find no significant effect on the h-index of the scientists who belong to a research cluster. Then, the academic novelty of the research performed is positively associated to the policy implementation, for both measures of novelty (average or maximum values), but this effect is once again not significant. Finally, we find that scientific research carried out by our treated researchers is not significantly more suitable with its incorporation afterwards in distinct disciplines, as captured by our two forward diversity index of knowledge diffusion across research fields.

Figure 3.17 – Fractional polynomial estimates of the number of citations (3-year window, and adjusted by # coauthors) obtained by yearly publications of the Core Bordeaux IdEx community (research clusters, in blue) and for their selected controls (in red).



Notes: The observed decrease in the number of (3-years) citations obtained by yearly publications in the second period can be related to the delayed referencing of recent articles in the WoS database. The last years are also biased downward considering the articles are published too recently to count forward citations in the full 3-year window.

Table 3.5 – Mean treatment effect estimated in the three years following the treatment assignment, and in the three next years, by difference-in-differences

Outcome variable	$[t+1\ ;\ t+3]$	$[t+4\ ;\ t+6]$
# Publications	0.0178	0.1522
	(0.228)	(0.344)
#Publications adjusted by co-authors	0.03	0.0513
	(0.051)	(0.071)
"Big hits"		
#Pub in top 10%	0.0592	0.0184
	(0.0871)	(0.122)
H-index	0.101	0.3026
	(0.205)	(0.446)
Normalized h-index	-0.027	-0.0491
	(0.054)	(0.064)
Novelty		
Average novelty	0.1375	0.0561
	(0.273)	(0.289)
Maximum novelty	0.1777	0.216
	(0.252)	(0.154)
Diffusion		
Simpson-Diversity	-0.001	-0.014
	(0.037)	(0.044)
Shannon	0.0333	-0.0043
	(0.091)	(0.106)

Notes: Fixed effects linear model with two periods, where the outcome measured before the treatment assignment is calculated on the period $[t-4\ ;\ t-2]$. Robust standard errors are reported into parenthesis and are clustered at the laboratory level. Each outcome variable is regressed separately on the interaction between the treatment dummy and time dummy, and Age and Age squared are included as independent variables in each model. The sample consists in 650 researchers (325 treated and 325 controls).

IV.2 Discussion

Our results suggest that IdEx policy did not significantly direct research carried out in the University of Bordeaux towards excellence, as measured by our selected indicators. Although research conducted just before the implementation of the policy may have been biased downward given that researchers devoted time to prepare proposals for this program, we are now able to track their research activity output during a sufficiently long time period that should allow us to observe the first consequences of the policy, if so.

However, these results have to be taken with caution for several reasons. First, it is possible that we do not cover all the dimensions that are affected by the policy. For instance, one important implication of the policy is that researchers affiliated to different laboratories are encouraged to carry out project together, while they may not have performed it otherwise. In that case, it is likely that targeted researchers tend to participate more in interdisciplinary projects, and may collaborate more with new coauthors. This, in turn, may lead to increase the skills and knowledge of these researchers and influence their future research agenda.

Then, part of our treated researchers may be relatively dissimilar from the potential controls in our raw sample, which implies that the optimal matching process we used lead to remove more than half of the treated scientists from the sample. Indeed, we previously removed all the universities and laboratories in France that are also involved in both IdEx or research cluster policy, which implies many qualitative researchers and laboratories. So our matching rather refers to an optimal subset matching (Rosenbaum 2012). The remaining treated researchers in the matched sample may thus not necessary be representative of the full sample of researchers targeted by the Idex policy in the University of Bordeaux. We perform a difference of means test between the treated in the matched sample and the full sample of treated (Table 3.6 in Appendix A). We observe significant mean differences, in terms of number of citations, h-index and h-index normalized. Researchers in the matched samples tend to be less cited than those in the full sample on average, and have lower h-index. They also are affiliated to smaller laboratories, where the most renowned member is significantly less cited compared to laboratories of the full sample. It then seems that researchers with better performances or affiliated to better laboratories have more difficulties to find a similar control during the matching process. Researchers (rather than professors) are also underrepresented in the matched sample compared to the full sample. So we cannot state that the results hold for our whole sample of core IdEx researchers, but they are rather related to a subsample which does not include part of the talented scientists.[14] Moreover, it should be remembered that researchers do not equally benefit from the policy, since the share of the funding dedicated to support research activity is mainly allocated to researcher in a competitive manner (through calls for proposals).

Finally, even though the matched sample is built upon several important individual and laboratory-level covariates that are finally strongly balanced, some observed information still remain imbalanced between the treated and control groups (such as the size of the laboratory). Moreover, the way the treatment is assigned to researcher implies a complex structure of the data. As many different information (observed at different levels) may be considered as potential confounders, it is still possible that we do not fully control for all relevant covariates.

For these reasons, we carry out the analysis again in the next subsection with the use of a redesigned control sample.

IV.3 Robustness Checks

Considering our matched sample failed to perfectly balance some laboratory-level covariates, we construct another control sample for which we set an additional constraint to fine balance the discipline of the laboratory. This constraint insures the balance of the marginal distribution of this variable between the treated and selected control groups. However, it also reduces the subset of control laboratories to be paired with our treated laboratories and finally produces a lower matched sample size of 520 researchers (260 treated paired to 260 controls units). Standardized difference in means are reported in Table 3.7 of Appendix B, in order to appraise the quality of the matching. The observed differences in terms of individual-level covariates in the raw sample have been substantially reduced with the matching, each of the covariates being strongly balanced again in the matched sample (with a remaining bias systematically lower than 10% of a standard deviation). When we turn to the investigation of the balance of the laboratory-level covariates, we see that the finely balanced match as expected removed the bias on the lab discipline. Thus, in the matched sample, the number of treated researchers affiliated to a laboratory of a particular discipline equals exactly the number of control researchers working in a laboratory of this same discipline. However, the priority given to this covariate may have worsen the balance of the other variables. While the covariate associated to the presence of a renowned researcher in the lab is strongly balanced (Max citations lab),

[14] Considering that the distribution of treated scientists into research fields are not always balanced between the two samples of treated, it may be that the variables which are based on citations (citations, h-index) are influenced by the share of researchers in highly productive disciplines (or on the contrary in research field with low publishing activity).

the categorical variables related to the size of the lab and the number of articles its members published are characterized by a remaining bias higher than 10% (but less than 20%).

Whereas this matched sample has been constructed in a slightly different way from the first control sample, prioritizing several variables that was not previously balanced, the results associated to the estimation of the policy impact do not change our conclusion drawn from the previous results (Table 3.8 in Appendix B). We again find that the effect of the policy on each of the outcome measures related to our three dimensions of research excellence, namely the scientific impact, novelty and knowledge diffusion, does not significantly differ from zero. This second matched sample therefore confirms the idea of no effect of the policy on researchers involved in the core IdEx policy of the University of Bordeaux.

V Conclusions

In this chapter we investigate the effects of implementing a policy targeting research sites and research clusters which aims at promoting research excellence of selected institutions. This policy is very ambitious as it aims to implement institutional change, at the University level as well as at the research groups level, to allow the treated reaching global scientific excellence.

Our study focuses on the University of Bordeaux, for which we are able to precisely identify the core community specifically targeted by the policy (who belongs to an IdEx and a research cluster). Moreover, we consider that enough time has elapsed since the implementation of the policy in 2011 to assess its first effects on our sample of targeted researchers. Considering the competitive nature of the selection of institutions and research clusters that benefit from the policy, we need to rely on the construction of a control group in order to estimate the causal effect of the assignment to treatment. For this purpose, we use a two-step optimal matching method developped by Zubizarreta & Keele (2017), that relies on the multilevel structure of the data. In our case, the matching is performed using information both at the individual and laboratory level, that allow us to control for individual characteristics along with the lab environment. With our matched sample, we then estimate the causal effect of the policy implementation by a difference-in-differences method on several outcome measures reflecting three dimensions of research excellence; the scientific impact, measured from the number of top cited papers or the h-index, novelty, which relates to the originality of research, and knowledge diffusion, as measured by the variety of discipline of the journals citing a paper.

We finally find that the ATT estimated on the outcome measures never differs significantly from zero, which indicates the absence of any effect of the policy implementation on our research excellence covariates for researchers involved in the core IdEx policy of the University of Bordeaux. This conclusion is still based on preliminary results that should be refined in order to investigate whether any confounding covariates may remain uncontrolled. Indeed, the balancing tests we performed on the matched sample signal that the main covariates specified for the matching are strongly balanced, but less important ones are still associated with higher level of imbalance (especially for some laboratory-level covariates). Moreover, we observe that the treated researchers included in the matched sample differ from the full sample of treated in terms of the mean number of citations, the h-index, the status and some characteristics of the laboratory. This gives some indications that the results may not be valid for the whole core Idex community, considering that part of the talented treated researchers are not considered in the analysis.

Moreover, these results should not be understood as evidencing the failure of the policy implementation (at least for the University of Bordeaux) since we only measure a limited number of dimensions of what can be defined as scientific excellence. In addition, it is very likely that this policy has consequences on research direction of the targeted researchers and influences to a certain extent who they collaborate with. Indeed, researchers are encouraged to develop projects with members of different laboratories, and are intended to carry out interdisciplinary projects.

Considering that one prerequisite of the policy is to gather different laboratories situated nearby to form research cluster, it would have also been possible to take into account the complementary nature of laboratories located in a specific geographic area, in order to represent the research cluster level. It would have implied to consider an additional level in the matching model, defined as the group of laboratories located in a research site. In that case, not only laboratories would have been paired, but it would also have allowed us to ensure that the control laboratories selected for a particular research cluster would all have been located in the same geographic area, just like main research clusters are. The disadvantage of this approach is that it may make the selection of similar researcher even harder.

For the purpose of this study, we have identified all the other laboratories involved in an IdEx or a research cluster in France. Since we only need controls affected by none of these policies, these labs have been removed from the database. However, in addition to the group of researchers not affected by each component of the policy, we identify sizable subsamples of researchers that fall into two other categories. In fact, we know laboratories involved in a research cluster but not in an IdEx, or on the contrary those not related to a research cluster but to an IdEx, which offers new opportunities to define the control group for future research. Thus, according to the way the control group is defined, we may be able to disentangle the effect of the two components of the policy. The comparison of the core IdEx researchers with controls who do not belong to an IdEx but are affiliated to a research cluster may allow us to identify the specific impact of IdEx, whereas controls belonging to an IdEx but not affiliated to a research cluster may allow us to appreciate the impact on the research cluster policy. Thereby, the complementarity of the two policies may be calculated by differentiating the joint effect from the sum of separate effects.

Appendix A. Difference of means test for the treated

Table 3.6 – Comparison of means between the treated researchers in the matched sample (matched) and the whole sample of treated (all)

Variable	Mean (matched)	Mean (all)	t	df	p-value
Age	42.6	42.14	-0.88	1140	0.37
Articles	0.604	0.665	1.32	1140	0.19
Citations	3.36	4.39	2.22	1140	0.026
Status *	0.37	44.5	2.35	-	0.018
H-index	5.93	6.85	3.17	1140	0.002
Norm H-index	0.75	0.88	4.05	1140	<0.001
Disciplines:					
Human Sciences *	0.068	0.07	0.12	-	0.9
Social Sciences *	0.01	0.24	-1.92	-	0.054
Mathematics *	0.16	0.08	3.28	-	0.001
Physics *	0.05	0.08	-1.93	-	0.053
Universe Sciences *	0.04	0	3.9	-	<0.001
Chemistry *	0.145	0.178	-1.38	-	0.17
Engineering *	0.04	0.05	-0.99	-	0.32
Information and Com. *	0.1	0.17	-3.57	-	<0.001
Agronomics and Ecology *	0.06	0.07	-0.86	-	0.39
Life Sciences and Environment *	0.13	0.12	0.36	-	0.72
Microbiology *	0.01	0.02	-0.82	-	0.41
Neurology *	0.11	0.05	3.27	-	0.001
Physiology and Endocrinology *	0.07	0.07	-0.02	-	0.98
Public health *	0.006	0.009	-0.57	-	0.57
Articles lab	2.02	2.07	0.88	1140	0.38
Max citations lab	89.05	100.89	2.56	1140	0.01
Lab size	39.88	49.44	6.12	1140	<0.001

Notes: Two-samples t-test. t denotes the statistic t and df stands for the degrees of freedom.

**Two-samples test of proportions.*

Status dummy variable: Researcher is 1 and Professor is 0.

Variables are measured in the three years before the treatment implementation.

Appendix B. Robustness Check estimation

Table 3.7 – Standardized difference of means between treated and control groups, before and after the matching

Individual-level covariates	Stand.bias Before	Stand.bias After
Age	-0.273	-0.033
Status	0.407	0.049
Norm H-index	0.525	0.074
Articles	0.275	0.052
Citations	0.323	0.053
Scientific field:		
Humanities	-0.251	0.012
Social Sciences	-0.364	0.00
Mathematics	0.344	0.013
Physics	-0.008	-0.017
Universe Sciences	-0.015	0.00
Chemistry	0.175	0.00
Engineering	-0.212	0.00
Information and Com. Sc. & Tech.	-0.053	0.00
Agronomics & Ecology	-0.045	0.00
Life Sciences & Environment	0.169	0.00
Microbiology	-0.133	0.00
Neurology	0.353	0.00
Physiology & Endocrinology	-0.088	-0.014
Public health	-0.126	0.00
Laboratory-level	**Stand.bias Before**	**Stand.bias After**
Articles lab	-0.284	-0.122
Max citations lab	-0.293	-0.029
Lab size	-0.067	-0.165
Main discipline of the lab:		
Biology	0.30	0
Health	-0.15	0
Ecology	0.186	0
Chemistry	0.383	0
Physics	0.194	0
Universe Sciences	-0.067	0
Engineering	0.006	0
Mathematics	-0.096	0
Human and Social sciences	-0.598	0

Notes: The unmatched sample is composed of 18,847 researchers and professors nested in 1,686 different laboratories (817 treated in 60 labs and 18,030 controls in 1,626 labs). The matched sample consists in 520 researchers (260 treated and 260 controls) affiliated to 120 different laboratories (60 labs for treated units paired to 60 other labs with control units).

Table 3.8 – Mean treatment effect estimated in the three years following the treatment assignment, and in the three next years, by difference-in-differences

Outcome variable	$[t+1\ ;\ t+3]$	$[t+4\ ;\ t+6]$
# Publications	-0.1405	-0.23
	(0.276)	(0.366)
#Publications adjusted by co-authors	0.0215	-0.0157
	(0.076)	(0.096)
"Big hits"		
#Pub in top 10%	-0.0782	-0.1260
	(0.086)	(0.131)
H-index	0.0479	0.2432
	(0.268)	(0.557)
Normalized h-index	-0.018	-0.0564
	(0.064)	(0.077)
Novelty		
Average novelty	0.297	0.4415
	(0.24)	(0.278)
Maximum novelty	0.2363	0.3222
	(0.216)	(.235)
Diffusion		
Simpson-Diversity	-0.0028	-0.0187
	(0.043)	(0.053)
Shannon	0.009	-0.0349
	(0.104)	(0.125)

Notes: Fixed effects linear model with two periods, where the outcome measured before the treatment assignment is calculated on the period $[t-4\ ;\ t-2]$. Robust standard errors are reported into parenthesis and are clustered at the laboratory level. Each outcome variable is regressed separately on the interaction between the treatment dummy and time dummy, and Age and Age squared are included as independent variables in each model. The sample consists in 520 researchers (260 treated and 260 controls).

Conclusion

In this thesis, we evaluate the efficiency and ex-post impact of two of the main policies implemented in France in the 2000s and present the characteristics of their funding allocation mechanism.

In Chapter 1, we investigate whether funding agencies behaves differently towards applicants who perform research which tackles original problems or interdisciplinary research, that we both define as components of non conventional research. This question stems from a growing concern in the scientific community about the risk aversion of panel committees when they select projects to fund. If so, novel and interdisciplinary research, that are both likely to produce considerable knowledge advance but are also associated to more uncertain results, could be penalized by review committees. We suspect that if applicants who carried out novel research or interdisciplinary research in the recent past are systematically less likely to be awarded a grant by the agency, this evidences a downward bias in the evaluation of their projects by the review committee. We test this hypothesis using applications submitted to the French funding agency (ANR). We find that researchers who address more original questions or perform higher degree of interdisciplinarity are more likely to apply to the agency. Nevertheless, these same researchers are negatively associated with the likelihood of an application to be successful. They are penalized by the peer-review committees. Everything works such as researchers who carry out unconventional research overestimate their chances to be awarded a grant. Indeed, they are more likely to apply, research quality and other factors being equals, but panel committees do not seem to appraise the projects in the same way the applicants do, since these researchers are less likely to be awarded a grant. In this context of information asymmetry, applicants know more their project than the committees, namely they have a private information which is not observable by the committees, which could explain this difference in the assessment of the project. In this way, it is likely that these researchers may in fact value more their work than other researchers do. When we differentiate the applications according to the type of program, we find that researchers who participate in higher level of interdisciplinary research are more likely to apply to both types of programs, whereas only the directed programs attract significantly more researchers who tackles very original problems. This result suggests that priority areas are well define by the agency for the directed programs. Nevertheless, unconventional research is still penalized during the grant process in the directed programs, whereas this only concerns interdisciplinary research in the non-directed programs. Moreover, other results suggest that the composition of the team is also correlated with the decision to fund a project. Project teams that are more multidisciplinary or more diverse in terms of cognitive differences are negatively associated with the probability to be awarded a grant. Projects are assessed by a mono-disciplinary review committee in the non-directed program, which may not be able to evaluate accurately interdisciplinary projects. However, panel committees are multi-disciplinary in the directed program.

In Chapter 2, we investigate the impact of obtaining an ANR project-based funding on grantees' ex-post scientific performances. We make use of the different program designs and recipients' characteristics in order to examine whether particular categories are relatively more affected by receiving a grant. The treatment effect is estimated using a group of controls weighted according to the estimation of the propensity score, which is defined as the probability for an applicant to be awarded a grant conditional on its observable attributes. Concerning the global efficiency of project-based funding, we find that obtaining a grant is associated with a substantial and significant increase in the number of citations received. Our results also suggest a positive and significant effect on the scope of collaboration network, both in terms of the size of the collaboration network and on the turn over of collaborators. In addition, we find a positive effect on the number of international collaborations. However, we do not find any evidence that obtaining a grant affects the novelty of the research questions tackled. Program design provides substantial variations in impact. Indeed, our results suggest that non directed programs are associated to a much larger impact than directed programs. Nonetheless the impact on directed programs is positive and significant, but it is rather small. Although we observe that directed programs succeed in attracting researchers who perform more novel research than applicants to non directed programs, we do not find evidence of any significant effect on the novelty of the research problems addressed ex-post by grantees. Moreover, we also find that young researchers benefit from a higher impact than older grantees.

In Chapter 3, we investigate the effectiveness of the IdEx program implementation. Our study focus on the case of the University of Bordeaux, and more precisely on researchers belonging to research clusters of excellence, who form the core target of Idex policy. Our goal is to examine whether this policy impacts ex-post research excellence of these scientists, which is defined in terms of scientific impact, novelty, and knowledge diffusion. In order to estimate the causal effect of this policy, we build a control group of French researchers who are similar to our core Bordeaux IdEx community but are not affected by the policy. The selection of the controls is based on individual attributes and scientific production measures, along with characteristics of the laboratory environment, which potentially confound the treatment effect. Our results suggest no significant impact of the policy on our research excellence measures for researchers involved in the core IdEx policy of the University of Bordeaux. However, these preliminary results should be viewed with caution for several reasons. First, although the covariates defined as the main potential confounders are strongly balanced between the control and treatment groups, inter-group differences remain for several other observable covariates. Secondly, our matching approach allows us only to keep half of the treated researchers, given that the more talented treated researchers do not manage to find close enough controls and are finally removed from the analysis. We also observe that the remaining core IdEx subsample differs significantly from the whole sample of treated researchers, so that our results can not be generalized to researchers belonging to a research cluster of the University of Bordeaux. Finally, even if there is in fact no significant effect on our research excellence indicators, we could not assert that the Idex policy is inefficient. Indeed, it is very likely that the way funding is allocated within clusters prompt researchers to develop interdisciplinary research projects and to collaborate with scientists they were not used to work with before.

We can draw lessons from this thesis. First, as the design of the programs (along with the way funding is allocated) is not neutral, it may affect who applies and who is finally funded, which in turn influences program efficiency. Indeed, we saw In Chapter 1 that different program designs attract and fund unlike applicants, which lead to different magnitude of impact according to the type of program In Chapter 2. Then, one of the missions of the French funding agency (ANR) is to promote original research projects and although we cannot observe directly the originality of the problem tackled by the project, we do not find evidence that this objective is achieved, neither in the selection of project (Chapter 1) nor in the ex-post impact on novelty (Chapter 2). This is an important concern often raised by funding agencies that would be interesting to examine more carefully.

Against this background, several other investigations could be conducted. First, considering we also identified the research laboratories involved in other research clusters and the institutions which were also selected by the IdEx program, this could represent new potential control groups with whom to compare the core Bordeaux IdEx community. It would allow for instance to separate the research cluster effect from the IdEx effect. The comparison between core Bordeaux Idex researchers and members of another research cluster who do not belong to an IdEx would bring the IdEx effect out, when the comparison with researchers who are affiliated to an Idex but not to a research cluster would emphasize the research cluster effect.

An interesting question would also focus on the relationship between these different types of funding, in order to study whether they are substitute or complementary. For instance, we could investigate whether belonging to a research cluster crowds-out or on the contrary crowds-in other project-based funding (such as ANR or European Research Council (ERC)). On one hand, considering project-based funding allocated within research clusters is probably associated with a higher success rate than those of ANR or ERC, research clusters members may not require to apply to alternative funding sources. On the other hand, it is also likely that talented teams who obtain grants from research clusters are also more likely to get complementary fundings from agencies.

We could also investigate whether researchers who are awarded a competitive grant benefit more from subsequent rewards, in order to test the principle of cumulative advantage (Merton 1968). For instance, we could use a duration model to examine whether funded researchers access more quickly subsequent competitive grants than not granted researchers, or whether they receive more rapidly promotion. Moreover, as we find a larger impact for young funded scientists in Chapter 2, it could be interesting to study whether obtaining a grant in the early career stages shifts lastingly the trajectory of the researcher.

Bibliography

Abadie, A. (2005), 'Semiparametric difference-in-differences estimators', *The Review of Economic Studies* **72**(1), 1–19.

Acemoglu, D., Johnson, S. & Robinson, J. A. (2012), 'The colonial origins of comparative development: An empirical investigation: Reply', *American Economic Review* **102**(6), 3077–3110.

Aghion, P., Dewatripont, M., Hoxby, C., Mas-Colell, A. & Sapir, A. (2010), 'The governance and performance of universities: evidence from europe and the us', *Economic Policy* **25**(61), 7–59.

Aghion, P., Dewatripont, M. & Stein, J. C. (2008), 'Academic freedom, private-sector focus, and the process of innovation', *The RAND Journal of Economics* **39**(3), 617–635.

Alberts, B. (2010), 'Overbuilding research capacity', *Science* **329**(5997), 1257.

ANR (2005), 'Agence nationale de la recherche, rapport annuel 2005', *accessible at "http://www.agence-nationale-recherche.fr/informations/documents/rapports-d-activite/"* .

Arora, A., David, P. A. & Gambardella, A. (2000), Reputation and competence in publicly funded science: estimating the effects on research group productivity, *in* 'The Economics and Econometrics of Innovation', Springer, pp. 141–176.

Arora, A. & Gambardella, A. (2005), 'The impact of nsf support for basic research in economics', *Annales d'Economie et de Statistique* pp. 91–117.

Arpino, B. & Mealli, F. (2011), 'The specification of the propensity score in multilevel observational studies', *Computational Statistics & Data Analysis* **55**(4), 1770–1780.

Arrow, K. J. (1972), Economic welfare and the allocation of resources for invention, *in* 'Readings in Industrial Economics', Springer, pp. 219–236.

Austin, P. C. (2009), 'Using the standardized difference to compare the prevalence of a binary variable between two groups in observational research', *Communications in Statistics-Simulation and Computation* **38**(6), 1228–1234.

Austin, P. C. (2011), 'An introduction to propensity score methods for reducing the effects of confounding in observational studies', *Multivariate behavioral research* **46**(3), 399–424.

Azoulay, P., Graff Zivin, J. S. & Manso, G. (2011), 'Incentives and creativity: evidence from the academic life sciences', *The RAND Journal of Economics* **42**(3), 527–554.

Banal-Estañol, A., Macho-Stadler, I. & Pérez-Castrillo, D. (2018), 'Team diversity evaluation by research grant agencies: Funding the seeds of radical innovation in academia?', *Working paper* .

Becker, S. O., Ichino, A. et al. (2002), 'Estimation of average treatment effects based on propensity scores', *The stata journal* **2**(4), 358–377.

Benavente, J. M., Crespi, G., Garone, L. F. & Maffioli, A. (2012), 'The impact of national research funds: A regression discontinuity approach to the chilean fondecyt', *Research Policy* **41**(8), 1461–1475.

Bertrand, M., Duflo, E. & Mullainathan, S. (2004), 'How much should we trust differences-in-differences estimates?', *The Quarterly journal of economics* **119**(1), 249–275.

Board, N. S. (2016), *Science & Engineering Indicators 2016*, National Science Foundation.

Bornmann, L., Leydesdorff, L. & Van den Besselaar, P. (2010), 'A meta-evaluation of scientific research proposals: Different ways of comparing rejected to awarded applications', *Journal of Informetrics* **4**(3), 211–220.

Bornmann, L., Wagner, C. & Leydesdorff, L. (2015), 'BRICS countries and scientific excellence: A bibliometric analysis of most frequently cited papers', *J. Assn. Inf. Sci. Tec.* **66**(7), 1507–1513.

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

Braben, D. W. (2004), *Pioneering research: A risk worth taking*, John Wiley & Sons.

Carayol, N. & Dalle, J.-M. (2007), 'Sequential problem choice and the reward system in open science', *Structural Change and Economic Dynamics* **18**(2), 167–191.

Carayol, N., Lahatte, A. & Llopis, O. (2018), 'The right job and the job right: Novelty, impact and journal stratification in science', *mimeo* .

Carayol, N. & Matt, M. (2006), 'Individual and collective determinants of academic scientists productivity', *Information Economics and Policy* **18**(1), 55–72.

Carley, S. & Porter, A. L. (2012), 'A forward diversity index', *Scientometrics* **90**(2), 407–427.

Carter, G. M., Winkler, J. D. & Biddle-Zehnder, A. K. (1987), *An evaluation of the NIH research career development award*, Rand Corporation.

Chubin, D. E., Hackett, E. J. & Hackett, E. J. (1990), *Peerless science: Peer review and US science policy*, Suny Press.

Chudnovsky, D., López, A., Rossi, M. A. & Ubfal, D. (2008), 'Money for science? the impact of research grants on academic output', *Fiscal Studies* **29**(1), 75–87.

Cochran, W. G. & Rubin, D. B. (1973), 'Controlling bias in observational studies: A review', *Sankhyā: The Indian Journal of Statistics, Series A* pp. 417–446.

Cole, S. & Simon, G. A. (1981), 'Chance and consensus in peer review', *Science* **214**(4523), 881–886.

Dasgupta, P. & David, P. A. (1994), 'Toward a new economics of science', *Research policy* **23**(5), 487–521.

Day, T. E. (2015), 'The big consequences of small biases: A simulation of peer review', *Research Policy* **44**(6), 1266–1270.

Dehejia, R. H. & Wahba, S. (1999), 'Causal effects in nonexperimental studies: Reevaluating the evaluation of training programs', *Journal of the American statistical Association* **94**(448), 1053–1062.

Dehejia, R. H. & Wahba, S. (2002), 'Propensity score-matching methods for nonexperimental causal studies', *Review of Economics and statistics* **84**(1), 151–161.

Egghe, L. & Rao, I. R. (2008), 'Study of different h-indices for groups of authors', *Journal of the American Society for Information Science and Technology* **59**(8), 1276–1281.

ERC (2017), 'European research council, annual report of the erc activities and achievements in 2017'.

Fang, F. C., Bowen, A. & Casadevall, A. (2016), 'Nih peer review percentile scores are poorly predictive of grant productivity', *Elife* **5**, e13323.

Fang, F. C. & Casadevall, A. (2016), 'Grant funding: Playing the odds', *Science* **352**(6282), 158–158.

Fleming, L. (2001), 'Recombinant uncertainty in technological search', *Management science* **47**(1), 117–132.

Frölich, M. (2004), 'Finite-sample properties of propensity-score matching and weighting estimators', *Review of Economics and Statistics* **86**(1), 77–90.

Geuna, A. (2001), 'The changing rationale for european university research funding: are there negative unintended consequences?', *Journal of economic issues* **35**(3), 607–632.

Geuna, A. & Martin, B. R. (2003), 'University research evaluation and funding: An international comparison', *Minerva* **41**(4), 277–304.

Geuna, A. & Rossi, F. (2015), *The university and the economy: pathways to growth and economic development*, Edward Elgar Publishing.

Ginther, D. K., Schaffer, W. T., Schnell, J., Masimore, B., Liu, F., Haak, L. L. & Kington, R. (2011), 'Race, ethnicity, and nih research awards', *Science* **333**(6045), 1015–1019.

González-Alcaide, G., Castellano Gómez, M., Valderrama-Zurián, J. C. & Aleixandre-Benavent, R. (2008), 'Literatura científica de autores españoles sobre análisis de citas y factor de impacto en biomedicina (1981-2005)'.

Guimera, R., Uzzi, B., Spiro, J. & Amaral, L. A. N. (2005), 'Team assembly mechanisms determine collaboration network structure and team performance', *Science* **308**(5722), 697–702.

Gush, J., Jaffe, A., Larsen, V. & Laws, A. (2018), 'The effect of public funding on research output: the new zealand marsden fund', *New Zealand Economic Papers* **52**(2), 227–248.

Hansen, B. B. & Bowers, J. (2009), 'Attributing effects to a cluster-randomized get-out-the-vote campaign', *Journal of the American Statistical Association* **104**(487), 873–885.

Heckman, J. J. (1976), 'The common structure of statistical models of truncation, sample selection and limited dependent variables and a simple estimator for such models', pp. 475–492.

Heckman, J. J. (1979), 'Sample selection bias as a specification error', *Econometrica* **47**(1), 153–161.

Heckman, J. J., Ichimura, H. & Todd, P. E. (1997), 'Matching as an econometric evaluation estimator: Evidence from evaluating a job training programme', *The review of economic studies* **64**(4), 605–654.

Hicks, D., Wouters, P., Waltman, L., Rijcke, S. d. & Rafols, I. (2015), 'The leiden manifesto for research metrics', *Nature* **520**(7548), 429–431.

Hirano, K. & Imbens, G. W. (2001), 'Estimation of causal effects using propensity score weighting: An application to data on right heart catheterization', *Health Services and Outcomes research methodology* **2**(3-4), 259–278.

Hirano, K., Imbens, G. W. & Ridder, G. (2003), 'Efficient estimation of average treatment effects using the estimated propensity score', *Econometrica* **71**(4), 1161–1189.

Hirsch, J. E. (2005), 'An index to quantify an individual's scientific research output', *Proceedings of the National academy of Sciences* **102**(46), 16569–16572.

Imbens, G. W. (2004), 'Nonparametric estimation of average treatment effects under exogeneity: A review', *Review of Economics and statistics* **86**(1), 4–29.

Imbens, G. W. & Wooldridge, J. M. (2009), 'Recent developments in the econometrics of program evaluation', *Journal of economic literature* **47**(1), 5–86.

Ioannidis, J., Boyack, K. W., Small, H., Sorensen, A. A. & Klavans, R. (2014), 'Bibliometrics: Is your most cited work your best?', *Nature News* **514**(7524), 561.

Ioannidis, J. P. (2011), 'More time for research: fund people not projects', *Nature* **477**(7366), 529.

Jacob, B. A. & Lefgren, L. (2011), 'The impact of research grant funding on scientific productivity', *Journal of public economics* **95**(9-10), 1168–1177.

Jacob, F. (1977), 'Evolution and tinkering', *Science* **196**(4295), 1161–1166.

Jones, B. F., Wuchty, S. & Uzzi, B. (2008), 'Multi-university research teams: Shifting impact, geography, and stratification in science', *science* **322**(5905), 1259–1262.

Kim, J. & Seltzer, M. (2007), 'Causal inference in multilevel settings in which selection processes vary across schools. cse technical report 708.', *National Center for Research on Evaluation, Standards, and Student Testing (CRESST)* .

Langfeldt, L., Benner, M., Sivertsen, G., Kristiansen, E. H., Aksnes, D. W., Borlaug, S. B., Hansen, H. F., Kallerud, E. & Pelkonen, A. (2015), 'Excellence and growth dynamics: A comparative study of the Matthew effect', *Sci. Public Policy* **42**(5), 661–675.

Laudel, G. (2006), 'Conclave in the tower of babel: how peers review interdisciplinary research proposals', *Research Evaluation* **15**(1), 57–68.

Lee, Y.-N., Walsh, J. P. & Wang, J. (2015), 'Creativity in scientific teams: Unpacking novelty and impact', *Research Policy* **44**(3), 684–697.

Lehman, H. C. (1953), *Age and achievement*, Princeton University Press, Princeton.

Levin, S. G. & Stephan, P. E. (1991), 'Research productivity over the life cycle: Evidence for academic scientists', *The American Economic Review* pp. 114–132.

Li, D. (2017), 'Expertise versus bias in evaluation: Evidence from the nih', *American Economic Journal: Applied Economics* **9**(2), 60–92.

Li, D. & Agha, L. (2015), 'Big names or big ideas: Do peer-review panels select the best science proposals?', *Science* **348**(6233), 434–438.

Li, F., Zaslavsky, A. M. & Landrum, M. B. (2013), 'Propensity score weighting with multilevel data', *Statistics in medicine* **32**(19), 3373–3387.

Li, L.-L., Ding, G., Feng, N., Wang, M.-H. & Ho, Y.-S. (2009), 'Global stem cell research trend: Bibliometric analysis as a tool for mapping of trends from 1991 to 2006', *Scientometrics* **80**(1), 39–58.

Mallard, G., Lamont, M. & Guetzkow, J. (2009), 'Fairness as appropriateness: Negotiating epistemological differences in peer review', *Science, Technology, & Human Values* **34**(5), 573–606.

Marsh, H. W., Jayasinghe, U. W. & Bond, N. W. (2008), 'Improving the peer-review process for grant applications: reliability, validity, bias, and generalizability.', *American psychologist* **63**(3), 160.

McDowell, J. M. (1982), 'Obsolescence of knowledge and career publication profiles: Some evidence of differences among fields in costs of interrupted careers', *The American Economic Review* **72**(4), 752–768.

Merton, R. K. (1957), 'Priorities in scientific discovery: a chapter in the sociology of science', *American sociological review* **22**(6), 635–659.

Merton, R. K. (1968), 'The matthew effect in science: The reward and communication systems of science are considered', *Science* **159**(3810), 56–63.

Merton, R. K. (1988), 'The matthew effect in science, ii: Cumulative advantage and the symbolism of intellectual property', *isis* **79**(4), 606–623.

Moed, H. F. (2008), 'UK research assessment exercises: Informed judgments on research quality or quantity?', *Scientometrics* **74**(1), 153–161.

Molinari, A. & Molinari, J.-F. (2008), 'Mathematical aspects of a new criterion for ranking scientific institutions based on the h-index', *Scientometrics* **75**(2), 339–356.

Möller, T., Schmidt, M. & Hornbostel, S. (2016), 'Assessing the effects of the german excellence initiative with bibliometric methods', **109**(3), 2217–2239.

Musselin, C. (2017), *La grande course des universités*, Presses de Sciences Po.

Nelson, R. R. (1959), 'The simple economics of basic scientific research', *Journal of political economy* **67**(3), 297–306.

Nicholson, J. M. & Ioannidis, J. P. (2012), 'Research grants: Conform and be funded', *Nature* **492**(7427), 34.

Park, H., Lee, J. & Kim, B.-C. (2015), 'Project selection in nih: A natural experiment from arra', *Research Policy* **44**(6), 1145–1159.

Petsko, G. (2012), 'Goodbye, columbus', *Genome biology* **13**(5), 155.

Pimentel, S. D., Page, L. C. & Keele, L. (2018), 'An overview of optimal multilevel matching using network flows with the matchmulti package in r'.

Post, C., De Lia, E., DiTomaso, N., Tirpak, T. M. & Borwankar, R. (2009), 'Capitalizing on thought diversity for innovation', *Research-Technology Management* **52**(6), 14–25.

Rafols, I. & Meyer, M. (2009), 'Diversity and network coherence as indicators of interdisciplinarity: case studies in bionanoscience', *Scientometrics* **82**(2), 263–287.

Reijnhoudt, L., Costas, R., Noyons, E., Börner, K. & Scharnhorst, A. (2014), ''Seed + expand': a general methodology for detecting publication oeuvres of individual researchers', *Scientometrics* **101**(2), 1403–1417.

Robins, J. M., Hernan, M. A. & Brumback, B. (2000), 'Marginal structural models and causal inference in epidemiology', *Epidemiology* **11**(5).

Rodrik, D. (2000), 'Institutions for high-quality growth: what they are and how to acquire them', *Studies in comparative international development* **35**(3), 3–31.

Rosenbaum, P. R. (1989), 'Optimal matching for observational studies', *Journal of the American Statistical Association* **84**(408), 1024–1032.

Rosenbaum, P. R. (2010), *Design of Observational studies*, New York: Springer-Verlag.

Rosenbaum, P. R. (2012), 'Optimal matching of an optimally chosen subset in observational studies', *Journal of Computational and Graphical Statistics* **21**(1), 57–71.

Rosenbaum, P. R., Ross, R. N. & Silber, J. H. (2007), 'Minimum distance matched sampling with fine balance in an observational study of treatment for ovarian cancer', *Journal of the American Statistical Association* **102**(477), 75–83.

Rosenbaum, P. R. & Rubin, D. B. (1983), 'The central role of the propensity score in observational studies for causal effects', *Biometrika* **70**(1), 41–55.

Rosenbaum, P. R. & Rubin, D. B. (1984), 'Reducing bias in observational studies using subclassification on the propensity score', *Journal of the American statistical Association* **79**(387), 516–524.

Rosenbaum, P. R. & Rubin, D. B. (1985), 'Constructing a control group using multivariate matched sampling methods that incorporate the propensity score', *The American Statistician* **39**(1), 33–38.

Rubin, D. B. (1974), 'Estimating causal effects of treatments in randomized and nonrandomized studies.', *Journal of educational Psychology* **66**(5), 688.

Schmoch, U. & Schubert, T. (2009), 'Sustainability of incentives for excellent research - the german case', *Scientometrics* **81**(1), 195–218.

Stephan, P. E. (1996), 'The Economics of Science', *Journal of Economic Literature* **34**(3), 1199–1235.

Stephan, P. E. (2012), *How economics shapes science*, Harvard University Press Cambridge, MA.

Stephan, P. E., Veugelers, R. & Wang, J. (2017), 'Blinkered by bibliometrics', *Nature* **544**(7651), 411–412.

Trapido, D. (2015), 'How novelty in knowledge earns recognition: The role of consistent identities', *Research Policy* **44**(8), 1488–1500.

Travis, G. & Collins, H. M. (1991), 'New light on old boys: Cognitive and institutional particularism in the peer review system', *Science, Technology, & Human Values* **16**(3), 322–341.

Uzzi, B., Mukherjee, S., Stringer, M. & Jones, B. (2013), 'Atypical combinations and scientific impact', *Science* **342**(6157), 468–472.

Van Raan, A. F. (2006), 'Comparison of the hirsch-index with standard bibliometric indicators and with peer judgment for 147 chemistry research groups', *scientometrics* **67**(3), 491–502.

van Steen, J. & Eijffinger, M. (1998), 'Evaluation practices of scientific research in the netherlands', *Research Evaluation* **7**(2), 113–122.

Wagner, C. S., Roessner, J. D., Bobb, K., Klein, J. T., Boyack, K. W., Keyton, J., Rafols, I. & Börner, K. (2011), 'Approaches to understanding and measuring interdisciplinary scientific research (idr): A review of the literature', *Journal of informetrics* **5**(1), 14–26.

Waltman, L. & Schreiber, M. (2013), 'On the calculation of percentile-based bibliometric indicators', *Journal of the American Society for Information Science and Technology* **64**(2), 372–379.

Wang, J., Lee, Y.-N. & Walsh, J. (2018), 'Funding model and creativity in science: Competitive versus block funding and status contingency effects', *Research Policy* **47**(6), 1070–1083.

Wang, J., Veugelers, R. & Stephan, P. E. (2017), 'Bias against novelty in science: A cautionary tale for users of bibliometric indicators', *Research Policy* **46**(8), 1416–1436.

Wessely, S. (1998), 'Peer review of grant applications: what do we know?', *The lancet* **352**(9124), 301–305.

Wuchty, S., Jones, B. F. & Uzzi, B. (2007), 'The increasing dominance of teams in production of knowledge', *Science* **316**(5827), 1036–1039.

Zhang, J., Yu, Q., Zheng, F., Long, C., Lu, Z. & Duan, Z. (2016), 'Comparing keywords plus of WOS and author keywords: A case study of patient adherence research: Comparing Keywords Plus of WOS and Author Keywords', *Journal of the Association for Information Science and Technology* **67**(4), 967–972.

Zubizarreta, J. R. & Keele, L. (2017), 'Optimal multilevel matching in clustered observational studies: A case study of the effectiveness of private schools under a large-scale voucher system', *Journal of the American Statistical Association* **112**(518), 547–560.

Zubizarreta, J. R., Paredes, R. D., Rosenbaum, P. R. et al. (2014), 'Matching for balance, pairing for heterogeneity in an observational study of the effectiveness of for-profit and not-for-profit high schools in chile', *The Annals of Applied Statistics* **8**(1), 204–231.

Zuckerman, H. & Merton, R. K. (1972), Age, aging, and age structure in science, *in* 'A Sociology of Age Stratification: Aging and Society', Russel Sage foundation, New York, Reprinted in: Storer, N.W. (Ed.), 1973. The Sociology of Science: Collected Papers of R.K. Merton, Chicago University, Chicago Press, 497-559.

List of Tables

List of Figures

Contents