DISCUSSION PAPER SERIES

IZA DP No. 9594

Connections in Scientific Committees and Applicants' Self-Selection: Evidence from a Natural Randomized Experiment

Manuel Bagues
Mauro Sylos-Labini
Natalia Zinovyeva

December 2015

Forschungsinstitut
zur Zukunft der Arbeit
Institute for the Study
of Labor

Connections in Scientific Committees and Applicants' Self-Selection: Evidence from a Natural Randomized Experiment

Manuel Bagues
Aalto University and IZA

Mauro Sylos-Labini
University of Pisa

Natalia Zinovyeva
Aalto University

Discussion Paper No. 9594
December 2015

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

IZA Discussion Paper No. 9594
December 2015

ABSTRACT

Connections in Scientific Committees and Applicants' Self-Selection: Evidence from a Natural Randomized Experiment*

We examine how the presence of connections in scientific committees affects researchers' decision to apply and their chances of success. We exploit evidence from Italian academia, where in order to be promoted to an associate or full professorship, researchers are firstly required to qualify in a national evaluation process. Prospective candidates are significantly less likely to apply when the committee includes, through luck of the draw, a colleague or a co-author. This pattern is driven mainly by researchers with a weak research profile. At the same time, information from 300,000 individual evaluation reports shows that applicants tend to receive more favorable evaluations from connected evaluators. Overall, this evidence is consistent with both the existence of a bias in favor of connected candidates and with academic connections reducing information asymmetries. Our study shows that connections are an important determinant of application decisions in academia and, more generally, it highlights the relevance of self-selection for empirical studies on discrimination.

JEL Classification: I23, M51, J45

Keywords: scientific evaluations, connections, self-selection

Corresponding author:

Manuel Bagues
Aalto University
Department of Economics
P.O. Box 21240
00076 Helsinki
Finland
E-mail: manuel.bagues@aalto.fi

* We would like to thank Mischa Drughov, Marta Martinez-Troya, Oskar Nordström Skans, Marko Terviö and participants at presentations at Universitat Autonoma de Barcelona, the Swedish Institute for Social Research, Helsinki Center of Economic Research, CERGE-EI Prague, the 2014 Trento Festival dell'Economia, Jyväskylä University, Annual Meeting of the Finnish Economic Association 2015, SOLE-EALE 2015, Labor Nordic Meeting 2015, "Brucchi Luchino" Workshop 2015 for their useful comments. All remaining errors are our own.

1 Introduction

It is well known that academic connections, such as those between co-authors, colleagues, advisors or mentors, are important for a successful professional career. In first place, they contribute to improve individual research productivity by providing ideas and intellectual support (Azoulay, Graff Zivin and Wang 2010, Oettl 2012). Moreover, researchers may also benefit from the presence in scientific committees of evaluators with whom there is some sort of prior connection (Combes, Linnemer and Visser 2008, Durante, Labartino and Perotti 2011, Li 2011, Perotti 2002, Sandström and Hällsten 2008), either because evaluators are positively biased towards connected candidates or because better information regarding the quality of candidates is more readily available (Brogaard, Engelberg and Parsons 2014, Laband and Piette 1994, Li 2011, Zinovyeva and Bagues 2015).

Beyond the direct impact of connections on evaluations, in this paper we examine whether the presence of a connection in a scientific committee affects researchers' decision to apply. The impact of connections on researchers' application decisions is theoretically ambiguous. If prospective candidates expect connected evaluators to be more favorable, this would increase the probability that researchers with a connection in the committee apply. On the other hand, connections may reduce information asymmetries between the candidate and the evaluator.[1] Evaluators may observe more accurately the quality of connected researchers, and potential applicants may also be better informed about the evaluation standards adopted by connected evaluators. Whether a reduction in information asymmetries dissuades or encourages connected researchers to apply depends on the quality of these same researchers.

We exploit the exceptional evidence provided by evaluations in Italian academia. Since 2012, in order to be promoted to associate and full professor, Italian researchers

[1] A number of studies have analyzed how information affects application decisions in other contexts. For instance, Hoxby and Avery (2013) study the application behavior of prospective college students in the US. They find that low-income high achievers apply to fewer selective colleges than their high-income counterparts with similar achievements, apparently because the latter are more likely to receive expert advice. Other authors have analyzed the impact of providing information to high school graduates about the labor market outcomes associated to different degrees (Hastings, Neilson and Zimmerman 2015, Pekkala-Kerr, Pekkarinen, Sarvimaki and Uusitalo 2015).

have had to qualify first in a scientific evaluation at a national level. Successful candidates can then apply for a promotion at the university level. Candidates who fail to qualify have to wait for two years before they can apply again. Given the cost of applying and the penalization faced by unsuccessful applicants, the decision to apply or not presents researchers with a dilemma. This set up has several features which are convenient for the purposes of analysis. In first place, it is wide-ranging. Its evaluations are conducted in every academic field and at two different stages of the career ladder (associate and full professorships). Second, researchers need to pre-register their application before the composition of the committee is known, independently of whether they finally apply or not. Thanks to this institutional feature, we are able to observe a list of prospective candidates. Third, committee members are randomly selected from a pool of eligible evaluators. This provides a credible and transparent empirical strategy. Finally, we observe the curriculum vitae of all potential candidates and evaluators, as well as evaluators' reports. We use this information to disentangle why connections may influence both the researchers' application decisions and these applications' chances of success.

Our database includes information on around 69,000 applications of researchers who pre-registered in 2012 for the first round of the national qualification evaluation. When the identity of committee members was announced, around 10,000 applications were withdrawn. The remaining 59,000 applications were evaluated by committee members and around 40% managed to qualify. We find that while the presence of connections in the committee decreases the probability that researchers apply, it increases their chances of success. The magnitude of these effects depends crucially on researchers' quality. In the case of researchers in the bottom tercile in terms of their research output, the presence of a connection in the committee decreases their application rate by 6 percentage points (p.p.) and it increases their (unconditional) chances of success by 3 p.p. On the other hand, researchers in the top tercile are as likely to apply when the committee includes a co-author or a colleague, and their probability of success is 5 p.p. larger.

To understand why candidates react to the presence of connections in committees, we examine the content of 300,000 individual reports, five per application. Connected candidates tend to receive more favorable evaluation reports. These reports tend also to be significantly longer, perhaps reflecting the availability of more accurate information. Furthermore, we also examine the application behavior and the performance of researchers who withdrew their application in the following round of evaluations, which took place in 2013. In this second round, which was evaluated by the same set of evaluators, connected researchers are more likely to apply and they are also more likely to qualify, relative to other researchers who withdrew their application in the first round. Overall, the evidence suggests that presence of a connection in the committee helps potential applicants whose research profiles are weak to submit their application at the best possible time and thus avoid costly mistakes.

Our paper contributes to the literature in several ways. Our results illustrate that academic networks provide access to information that helps to make better professional choices. The information provided by connections may be useful also for other important academic decisions such as applying for a grant, for a position, or selecting the outlet where a paper should be submitted, a process that is also costly and subject to uncertainty.[2]

Moreover, to the best of our knowledge, this paper provides the first evidence showing that the composition of scientific committees affects application decisions. This finding has important methodological implications for the empirical analysis of evaluation biases and discrimination. If candidates self-select into the application process on the basis of the identity of evaluators, this may bias in a non-trivial way studies that rely only on information about actual applicants. To deal with self-selection, it might be necessary to consider all prospective applicants, independently of whether they apply or not. The endogenous self-selection of candidates may be also relevant

[2]This informational feature of connections might partly explain the success of some mentoring programs. For instance, Blau et al. (2010) randomly selected some junior female economists from a set of volunteers and provided them with an intensive two-day mentoring on "research and publishing, getting grants, professional exposure, teaching, the tenure process and work-life balance." Five years after the intervention, participants in the workshop were 20 percentage points more likely to have a top-tier publication and have 2.7 more publications overall.

for the interpretation of audit and correspondence studies. In these studies, fictitious applicants look identical "on paper" except for some particular characteristic such as gender or race. As pointed out by Heckman and Siegelman (1993) and Neumark (2012), an evaluator's decision to select applicants from a certain group may reflect either taste discrimination or, if different groups of candidates are expected to differ in their distribution of quality in some unobserved dimension, it may reflect statistical discrimination. Our analysis suggests that, even if the two groups are identical in the overall population, they are likely to differ among applicants due to the self-selection of applicants.[3]

Finally, our study may also contribute to improve the design of scientific evaluations. Policy makers may want to consider more carefully whether prospective applicants should receive information about the identity of evaluators. For instance, in the context of the qualification exams that we study in this paper, allowing pre-registered candidates to withdraw their application once the committee composition is announced amplifies the benefits of connections, allowing some connected candidates with a weak research profile to withdraw their application and avoid a costly and time-consuming failure.

The paper is organized as follows. We start by proposing a simple model of application behavior that helps to clarify how connections in committees may affect candidates' decision to apply. In section 3, we explain the structure of the evaluation process. In section 4, we describe the data used in the empirical analysis and in section 5 we present our main findings. In section 6 we briefly summarize the findings and we discuss possible interpretations and implications.

[3]This might perhaps explain the results in Milkman, Akinola and Chugh (2015), who conduct an audit study in which fictional prospective students contact professors in order to discuss research opportunities prior to applying to a doctoral program. Faculty are significantly less responsive to students with a foreign-sounding name even if, by construction, their messages were otherwise identical. A possible explanation, within the framework of our study, is that employers prejudge native prospective students to be better informed about their fit and, as a result, they foresee that they will be positively selected among students who decide to contact the faculty.

2 Theoretical framework

We propose a simple conceptual framework to analyze how the presence of a connection in a committee may affect prospective candidates' decision to participate in an evaluation process. The model captures three relevant features. First, applications tend to involve some costs, either in the form of specific investments or opportunity costs. Given these costs, candidates need to weigh up their probabilities of success and carefully consider whether they should apply or not. Second, the outcome of the evaluation may depend on the identity of the evaluator. Evaluators may have a preference for certain areas of research or they may be biased in favor or against some candidates. Third, there might be relevant information asymmetries both on the evaluators' and on the researchers' side. While evaluators may observe imperfectly the quality of candidates, candidates may likewise not be perfectly informed about evaluators' standards or about their preferences.

According to the model, the impact of connections on application decisions is ambiguous. If evaluators are positively biased towards connected researchers, this would increase the likelihood that these researchers apply. The opposite would be true if they are negatively biased. Moreover, if connections convey information on evaluation standards to potential applicants or if they provide information to evaluators on the quality of candidates, the impact of connections on applications can be either positive or negative depending on the quality of connected researchers.

2.1 Set up

More formally, let us consider an individual i of quality q_i who has to decide whether to submit an application to evaluator j. The net gain of applying and qualifying is equal to G while the cost of applying and failing is equal to C, where both G and C are positive. The payoff for the individual if he does not apply is equal to zero. The payoff function of the candidate can be described as follows:

$$\text{Payoff}\,_{ij}^{cand} = \begin{cases} G, & \text{if candidate } i \text{ applies and qualifies,} \\ -C, & \text{if candidate } i \text{ applies and fails,} \\ 0, & \text{if candidate } i \text{ does not apply.} \end{cases}$$

If the candidate applies, the evaluator assesses the evaluator assesses the application and the payoff function is as follows:

$$\text{Payoff}\,_{ij}^{eval} = \begin{cases} q_i + B_{ij}, & \text{if candidate } i \text{ is granted a qualification,} \\ U_j, & \text{if candidate } i \text{ is not granted a qualification,} \end{cases}$$

where B_{ij} reflects the potential existence of subjective bias and U_j is the outside option of the evaluator or, equivalently, the threshold that the candidate needs to achieve in order to be granted a qualification.

There are two sources of uncertainty in the model. First, the potential applicant does not know precisely how large the threshold U_j is. He has some distributional prior information, $U_j \sim N(0,1)$ and, additionally, he receives a private signal about the actual draw of U_j:

$$z_{ij} = U_j + \epsilon_{ij}, \quad \epsilon_{ij} \sim N(0, \gamma_{ij}^2),$$

where γ_{ij}^2 reflects the degree of accuracy of the signal that the individual receives. The second source of uncertainty comes from the fact that the evaluator observes only imperfectly the true quality of the candidate. She knows the distribution of quality among prospective candidates, that for the sake of simplicity is $q_i \sim N(0,1)$, and she also receives a private signal about the actual quality of the candidate:

$$y_{ij} = q_i + \eta_{ij}, \quad \eta_{ij} \sim N(0, \sigma_{ij}^2),$$

where σ_{ij}^2 is the accuracy of the signal. While the candidate and the evaluator do not observe the signals received by each other, prior beliefs and the accuracy of the signals are common knowledge.

Let us derive the application decision of the prospective applicant by means of backward induction. First, consider the second stage, at which the evaluator decides whether to promote or fail the candidate. For simplicity, let us assume that the evaluator only takes into account the observed signal and the prior distributional information, and she does not try to infer the quality of the applicants based on their application decisions.[4] The evaluator promotes the candidate whenever his expected quality is higher than the outside option:

$$E(q_i + B_{ij}|y_{ij}) = \frac{y_{ij}}{1+\sigma_{ij}^2} + B_{ij} > U_j \Rightarrow \text{ promote candidate } i. \tag{1}$$

Now let us consider the first stage, at which the candidate decides whether to apply. The candidate forms a judgment about how his application will be perceived by the evaluator. This judgment takes both the candidate's own quality into account as well as the accuracy of the signal that the evaluator will observe:

$$E(q_i|y_{ij})|q_i \sim N\left(\frac{q_i}{1+\sigma_{ij}^2}, \frac{\sigma_{ij}^2}{(1+\sigma_{ij}^2)^2}\right).$$

At the same time, the candidate also forms a posterior distribution about the grading standards of the evaluator, based on the private signal that he receives:

$$U_j|z_{ij} \sim N\left(\frac{z_{ij}}{1+\gamma_{ij}^2}, \frac{\gamma_{ij}^2}{1+\gamma_{ij}^2}\right).$$

Given the decision rule of the evaluator in the second stage (equation (1)), the expected probability that the candidate will qualify is equal to:

$$Pr\left(E(q_i + B_{ij}|y_{ij}) > U_j|q_i, z_{ij}\right) = \Phi\left(\frac{\frac{q_i}{1+\sigma_{ij}^2} + B_{ij} - \frac{z_{ij}}{1+\gamma_{ij}^2}}{\sqrt{\frac{\sigma_{ij}^2}{(1+\sigma_{ij}^2)^2} + \frac{\gamma_{ij}^2}{1+\gamma_{ij}^2}}}\right), \tag{2}$$

where $\Phi(\cdot)$ is the cumulative density function of a standard normal distribution. Assuming risk neutrality, individual i will be willing to apply as long as, based on the available information, the expected net return from applying is positive. Candidate i

[4] There are two possible ways to interpret this simplifying assumption. Formally, we may think of a context where committee members evaluate researchers without knowing whether they are applying or not. Alternatively, we may consider naive evaluators, who are unaware of the fact that candidates' decision to apply may reveal information about their quality.

applies as long as:

$$\begin{aligned}
&Pr(E(q_i + B_{ij}|y_{ij}) > U_j|q_i, z_{ij}) * G - [1 - Pr(E(q_i + B_{ij}|y_{ij}) > U_j|q_i, z_{ij})] * C > 0 \\
&Pr(E(q_i + B_{ij}|y_{ij}) > U_j|q_i, z_{ij}) > \frac{C}{G + C}. \qquad (3)
\end{aligned}$$

Notice that the prospective candidate of quality q_i applies if he receives a sufficiently low signal about the evaluation threshold U_j. To see this, we can substitute the expression for the probability of success (2) in the application rule (3), rearrange the terms and express the application rule in the following form:

$$z_{ij} < z^*_{ij} \Rightarrow \text{candidate } i \text{ applies}, \qquad (4)$$

where $z^*_{ij} = \left[\frac{q_i}{1+\sigma^2_{ij}} + B_{ij} - \Phi^{-1}\left(\frac{C}{G+C}\right)\sqrt{\frac{\sigma^2_{ij}}{(1+\sigma^2_{ij})^2} + \frac{\gamma^2_{ij}}{1+\gamma^2_{ij}}}\right](1 + \gamma^2_{ij})$ and $\Phi^{-1}(\cdot)$ is the inverse cumulative density function.

Given this application rule, let us now analyze how the probability that a prospective candidate applies varies depending on his own quality q_i and on the evaluator's grading standards U_j:

$$Pr(z_{ij} < z^*_{ij}|U_j, q_i) = \Phi\left(\frac{\left[\frac{q_i}{1+\sigma^2_{ij}} + B_{ij} - \Phi^{-1}\left(\frac{C}{G+C}\right)\sqrt{\frac{\sigma^2_{ij}}{(1+\sigma^2_{ij})^2} + \frac{\gamma^2_{ij}}{1+\gamma^2_{ij}}}\right](1 + \gamma^2_{ij}) - U_j}{\gamma_{ij}}\right). \qquad (5)$$

2.2 Comparative statics

We use expression (5) to analyze the three channels through which connections might have an impact on application behavior: evaluation bias, lower uncertainty about the candidate's quality, and lower uncertainty about the evaluator's standards.

Case 1: connections and evaluation bias. First, let us consider the case when there is an evaluation bias (connections affect B_{ij}) but connections do not reduce information asymmetries (σ^2_{ij} and γ^2_{ij} are constant). Since $\Phi(\cdot)$ is a monotonically increasing function, the probability of applying increases in B_{ij} for all candidates. If connections do only involve a positive (negative) evaluation bias, we should observe an increase (decrease) in the probability that connected candidates apply.

Case 2: connections convey information on candidates. The situation is different if connections reduce information asymmetries. As we show below, depending on candidates' quality, connections in the committee might either encourage or discourage candidates from applying. Consider the possibility that connected candidates are better informed about evaluators' preferences (connections reduce γ_{ij}^2). For simplicity, let us assume that there is no evaluation bias ($B_{ij} = 0$), the evaluator can perfectly observe candidate quality ($\sigma_{ij}^2 = 0$), and $C = G$. The probability that the candidate applies is equal to:

$$Pr(z_{ij} < z_{ij}^*|U_j) = \Phi\left(\frac{q_i(1+\gamma_{ij}^2) - U_j}{\gamma_{ij}}\right).$$

The derivative of this expression with respect to γ_{ij} is:

$$\frac{\partial Pr(z_{ij} < z_{ij}^*|U_j)}{\partial \gamma_{ij}} = -\phi\left(\frac{q_i(1+\gamma_{ij}^2) - U_j}{\gamma_{ij}}\right)\frac{q_i(1-\gamma_{ij}^2) - U_j}{\gamma_{ij}^2},$$

where $\phi(\cdot)$ is the probability density function of a standard normal distribution. The sign of this derivative depends on the values of q_i, γ_{ij}^2 and U_j. A reduction in the uncertainty regarding the evaluation threshold would induce a relatively good candidate ($q_i > \frac{U_j}{1-\gamma_{ij}}$) to apply more and a relatively weak candidate ($q_i < \frac{U_j}{1-\gamma_{ij}}$) to apply less.

Case 3: connections convey information on evaluation standards. Consider now the case when evaluators observe more accurately the quality of connected candidates (connections reduce σ_{ij}^2). Again, for simplicity let us assume that the candidate can perfectly observe grading standards ($\gamma_{ij}^2 = 0$), there are no evaluation biases ($B_{ij} = 0$), and $C = G$. The probability that a prospective candidate applies is equal to:

$$Pr(z_{ij} < z_{ij}^*|U_j) = \begin{cases} 1, & \text{if } \frac{q_i}{1+\sigma_{ij}^2} - U_j > 0, \\ 0, & \text{if } \frac{q_i}{1+\sigma_{ij}^2} - U_j < 0. \end{cases}$$

The candidate would only apply if, given his quality, he expects that the evaluator will observe a high enough signal. A more precise signal would increase the candidate's willingness to apply if candidate quality is above average ($q_i > 0$). On the contrary, below average quality candidates are less likely to apply when the signal is more informative.

In sum, the nature of the connections might determine whether there is an increase or a decrease in the prospective candidate's willingness to apply. If connections in committees are mainly associated with a positive evaluation bias, we would expect that connected candidates are negatively selected into the application. However, if connections decrease information asymmetries between the candidate and the evaluator, the effect of connections on applications is ambiguous. Relatively weak candidates would be less likely to apply when they have a connection in the committee and, by doing so, they would avoid the cost of failure. On the contrary, candidates who excel in dimensions that are observed more accurately by connected evaluators would be more likely to apply.

3 Background

Most Italian universities are public and the recruitment of full and associate professors is regulated by national laws.[5] Before 2010, recruitment procedures were managed locally by each university. In 2010, a two-stage procedure similar to those already in place in other European countries was approved (e.g. France and Spain).[6] In the first stage, candidates to associate professor and full professor positions are required to qualify in a national-level evaluation known as the National Scientific Qualification (*Abilitazione Scientifica Nazionale*). Evaluations are conducted separately in 184 scientific fields designed by the Ministry of Education. A positive evaluation is valid for four years while a negative one implies a ban on participating in further national evaluations during the following two years. Qualified candidates can participate in the second stage, which is

[5]According to *OECD Education at a glance* (2013), in 2011 about 92% of students in tertiary education were enrolled in 66 public universities and the remaining 8% in 29 independent private institutions.

[6]Law number 240/2010, also known as "Gelmini reform" after the name of the minister of Education.

managed locally by each university. The introduction of a qualification exam at the national level was intended to reduce nepotism, which according to some authors was prevalent in the previous system (Durante et al. 2011).

3.1 The National Scientific Qualification

The first National Scientific Qualification was performed between 2012 and 2014.[7] The timeline of the process is described in Figure 1. The call for eligible evaluators was published in June 2012. The deadline for professors to volunteer to be an evaluator was August 28. Once the list of eligible evaluators was settled, the Ministry publicized their identities and their CVs. In the meantime, the call for candidates' applications was issued in July. Candidates had to pre-register online by November 20. The submission package included the CV and up to 20 selected publications. Researchers were able to apply to multiple fields and positions.

Once the application deadline was closed, committee members were selected by random draw. These lotteries were held between late November 2012 and February 2013. Following their appointment, and before the list of pre-registered applicants was known, each evaluation committee had to draft and to publish online a document describing the general criteria that would be used to grant positive evaluations.[8] At this point, pre-registered candidates could still withdraw their application. The deadline to withdraw the application expired two weeks after the committee composition had been decided and the committee had publicly announced the evaluation criteria. By the end of this period, evaluation committees were informed about the final list of candidates and the examination took place. Below we explain in more detail how committee members were selected and the evaluation process.

[7] A detailed description of the process is available at `http://abilitazione.miur.it/public/index.php?lang=eng`, retrieved on February 2014.

[8] For instance, in Econometrics the committee announced that "(i)n order to assess the scientific maturity of the candidates, the Committee will give prominent weight to the evaluation of their scientific publications, especially those published in top journals. The publications will be evaluated on the basis of their originality, innovativeness, methodological rigor, international reach and impact, and relevance for the field. In order to evaluate journal articles, the Committee may use the classification of journals provided by ANVUR and the bibliometric indicators provided by Web of Science and Scopus. The Committee may also use information regarding the impact of each individual publication and the total number of citations received by the candidate."

3.2 Selection of committees

The pool of eligible evaluators includes full professors in the corresponding field who have volunteered for the task and satisfy some minimum quality requirements. Math, engineering, and natural and life sciences require a research production which is above the median for full professors in the field and which is present in at least two of the following three dimensions: (i) the number of articles published in scientific journals covered by ISI Web of Science, (ii) the number of citations, (iii) and the H-index.[9] In the social sciences and the humanities, eligible evaluators are required to have a research production above the median in at least one of the following three dimensions: (i) the number of articles published in high quality scientific journals (in what follows, A-journals),[10] (ii) the overall number of articles published in any scientific journals and book chapters, and (iii) the number of published books.

Eligible evaluators may be based in Italy (hereafter 'Italian') and may also be affiliated to a university from an OECD country (hereafter 'international'). International and Italian eligible evaluators have to satisfy the same research requirements but their remuneration differs. While 'Italian' evaluators work *pro bono*, OECD evaluators receive €16,000 for their participation.

Evaluation committees include five members. Four members are randomly drawn from the pool of eligible Italian evaluators, under the constraint that no university can have more than one evaluator within the committee. The fifth member is typically selected from the pool of eligible international evaluators. Exceptionally, whenever the pool of international professors includes less than four professors, all five committee members are drawn from the pool of eligible evaluators based in Italy. Randomization is conducted in a way that leaves little room for manipulation. Eligible evaluators in each field are ordered alphabetically and are assigned a number according to their position. A sequence of numbers is then randomly selected. The same sequence is

[9] More precisely, this rule applies to Mathematics and IT, Physics, Chemistry, Earth Sciences, Biology, Medicine, Agricultural and Veterinary Sciences, Civil Engineering and Architecture (with the exception of Design, Architectural and Urban design, Drawing, Architectural Restoration, and Urban and Regional Planning), Industrial and Information Engineering, and Psychology.

[10] An evaluation agency and several scientific committees determined the set of high-quality journals in each field.

applied to select committee members in a number of different fields.

Evaluators are in charge for two years. If an evaluator resigns, a substitute evaluator is selected randomly from the corresponding group of eligible evaluators.

3.3 The evaluation

The evaluations are (officially) based only on candidates' CVs and publications. There are no oral or written tests or interviews. Committee members meet periodically to discuss their assessments and cast their votes. A positive assessment requires a qualified majority of four positive votes (out of five committee members).

Committees have full autonomy on the exact criteria to be used in the evaluation. Nonetheless, it is important to point out that an independent evaluation agency (ANVUR), appointed by the Ministry, collected and publicized information on the research productivity of all candidates in the previous ten years This productivity was first measured by the same three bibliometric indicators employed to select evaluators and it was then normalized by taking into account the amount of time passed since first publication and also the number of job interruptions (this last typically related to parental leave). The evaluation agency also used these bibliometric dimensions to provide the average research productivity of professors in those categories to which candidates might apply. Committees are not obliged, though encouraged, to use this information.

At the end of the process, committees provide each candidate with (i) the final outcome of the evaluation (pass or failure), (ii) a collective report explaining the criteria used by the committee and how they reached their final decision and (iii) five individual reports explaining each evaluators' position. Figure (2) provides a sample of an individual evaluation report.

4 Data

We consider all evaluations held within the first edition of the *National Scientific Qualification*.[11] The database includes examinations for associate and full professorships in 184 academic fields. We describe below the available information on (i) the pool of eligible and actual evaluators; (ii) the pool of pre-registered and actual applicants and (iii) the final outcome of the evaluation.

4.1 Evaluators

Around six thousand professors , all based in Italy, volunteered and qualified to be in the pool of eligible evaluators. The number of professors in the pool of eligible evaluators based abroad was slightly above one thousand. In the average field, the pool of eligible evaluators includes 32 Italian professors and eight international professors.

Table 1 provides some descriptive information on eligible evaluators. The share of women is low both among Italian eligible evaluators (20%) and international ones (12%). The average CV includes around 131 research outputs, mostly journal articles (73), book chapters (22), and conference proceedings (20). The average CV also includes 0.42 patents. As a proxy for the quality of journal articles, we have collected information on the quality of the journals in which they were published. In social sciences and humanities we use the official list of A-journals that was compiled by the evaluation agency. This list includes approximately 7,000 academic journals. Eligible evaluators have published on average 11 articles in A-journals. In sciences, we consider the *Article Influence Score* (AIS) of journals.[12] The AIS of the average publications of an eligible evaluator is 1.18.

Approximately 8% of Italian evaluators drawn in the initial lottery resigned and were replaced by other (randomly selected) eligible evaluators. The resignation rate

[11] We collected the CVs of candidates and evaluators and the final evaluations from the webpage of the Ministry of Education. To avoid problems with homonymity, we have excluded 14 candidates that had the same name and surname as other candidates within the same field and rank.

[12] This indicator is available for all publications in the Thomson Reuters Web of Knowledge. It is related to *Impact Factor*, but it takes into account the quality of the citing journals, the propensity to cite across journals and it excludes self-citations. The average journal is normalized to have AIS equal to one.

was slightly higher among international evaluators (10%).[13]

4.2 Applications

The National Scientific Qualification attracted 46,236 candidates and 69,020 applications. This is a relatively large number, accounting for around 61% of assistant professors and 60% of associate professors in Italy.[14] One third of candidates registered in several fields (e.g: qualification to full professorship in Political Economy and qualification to full professorship in Applied Economics) or in different categories of the same field (e.g.: qualification to full and associate professorships in Political Economy). In total there were approximately 375 applications per field.

In the upper panel of Table 2, columns 1 and 2 provide information on the characteristics of the initial set of applications. Columns 3 and 4 distinguish between candidates to a position of full and associate professor. As expected, in evaluation exams for a position of full professor applicants tend to be relatively older (49 vs. 43 years old) and are less likely to be female (31% vs. 41%). Applicants to full professorships are also more likely to hold a permanent position in an Italian university (74% vs. 47%). In roughly three fourths of the cases, researchers with a permanent position are applying to the same field where they officially hold this position.

The average CV has 16 pages and it reports 64 research outputs, mostly journal articles (37). It includes also some books (2), book chapters (7), conference proceedings (10), and patents (0.24). Not surprisingly, candidates to full professor positions have a relatively longer publication record: 89 vs. 53 publications. A typical paper is coauthored by six authors, with only 34% of papers being single authored. The candidate reports to be the first author in 22% of the occasions. In social sciences and humanities, the average candidate for a position of full professor has published six articles in A-journals; applicants to associate professorships have published about three articles

[13] In two fields where the international member of the committee resigned, the pool of international evaluators included originally just four members. In these two cases, given that the pool of remaining eligible evaluators was lower than four, the replacement was selected from the Italian pool.

[14] Source: Our own calculations using information from the Italian Ministry of Education on the identity of all assistant (*Ricercatori*) and associate professors (*Associati*) in Italy on December 31 2012.

less. In sciences, the average AIS of papers published by candidates for a position of full professor is around 1.31; it is similar for candidates to associate professorships. Compared to eligible evaluators (who are a selected sample of full professors), the publication record of applicants for a position of full professor is roughly 35% shorter.

As explained earlier, based on candidates' CVs, the evaluation agency of the Ministry of Education constructed and publicized detailed information on candidates' research production during the 10 previous years measured along three dimensions. This information, available only for candidates who did not withdraw their application (86%), is summarized in the lower panel of Table 2. Around 38% of candidates were above the median in each of the three dimensions. On the other end of the scale, 16% of candidates were below the median in every dimension.

We have also constructed a proxy for the timing of the application. We use the application code number, which reflects the ordering of application, and we normalize this variable uniformly between 0 and 1 for applicants within the same list. The timing of the application might perhaps be correlated to candidates' quality or with their self-confidence.

4.3 Evaluations

Approximately 14% of applications were withdrawn by applicants when the identity of evaluators and the general evaluation criteria were revealed. The remaining applications received an evaluation from the committee. Table 3 provides information on the outcome of the evaluation process. Out of the 59,150 applications that received an evaluation, 43% were successful. The success rate is slightly lower if we consider all applications, including those that were withdrawn (37%). Success is strongly correlated with candidates' observable research productivity. As shown in Figure 3, among actual candidates whose quality was below the median in every dimension, only 4% managed to succeed. On the contrary, 63% of candidates that excelled in all three dimensions qualified.

Each committee member writes an individual evaluation report for each application. Overall there are approximately 295,000 individual reports.[15] The average report includes around 176 words and provides a description of the research production of the candidate, some discussion about its quality and its fit with the field. It also indicates the evaluator's final assessment on whether the candidate deserves qualification. We have conducted a text analysis of these reports in order to identify the final assessment. On most occasions, the final assessment was decided unanimously by all five evaluators (86%). Over all, 45% of votes were favorable to the candidate and 55% were negative.

Those candidates who had withdrawn the application in the first round of evaluations had a chance to participate in the evaluations conducted the following year (this is the last evaluation round conducted so far). Around 37% of these candidates chose to reapply. We have collected information on the outcomes in the second round of evaluations.[16] Out of the group of those who had reapplied, 66% managed to qualify.

4.4 Connections

We consider two types of links between candidates and evaluators: co-authorships and affiliation to the same institution. In the National Scientific Qualification these links are not formally subject to a conflict of interest rule and, depending on the result of the random draw, candidates might be evaluated by a professor affiliated to the same university or by a co-author.[17]

These two links, colleagues and co-authors, may be associated to different features. Colleagues are in general expected to be close in social terms but not necessarily intellectually. They might have private information on candidates' contribution to professional service and, sometimes, they might be perhaps directly affected by the outcome of the evaluation. Co-authors tend to be close both in the social space and the ideas space.

[15]Due to a technical problem, we are missing information on evaluation reports of 202 applications.

[16]One committee had not published their evaluations for the second-round applicants as on September 15, 2015.

[17]Officially, only kinship relationships are subject to a restriction. Nonetheless, committees might autonomously decide to self-impose their own additional restrictions. According to our analysis of the evaluation reports, evaluators voluntarily abstained in the presence of a colleague or a co-author in only three fields (out of a total 184). These three fields are Ecology (sector 05/C1), Pediatrics (06/G1) and Management (13/B2). As a result, 84 candidates in these fields received only four evaluation reports.

In practice, approximately 12% of candidates are assigned to a committee that includes a colleague and around 7% to a committee including a co-author.[18] In about a third of the cases the evaluators who have co-authored with the candidate also belong to the same university.

Candidates with a connection in the evaluation committee, either a colleague or co-author, tend to have a significantly better research profile relatively to the rest of candidates (Table 2, columns 5-7). Connected candidates excel both in terms of quantity and quality of research, probably reflecting that eligible evaluators are positively selected combined with the existence of assortative matching in co-authorship decisions and affiliations.

5 Empirical analysis

Using the evidence provided by scientific evaluations in Italy, we study the role of two specific types of academic connections: colleagues and co-authors. We estimate their causal impact upon researchers' application decisions and we also examine which of the two mechanisms, bias or information, is consistent with the evidence. According to the conceptual framework presented in section 2, if evaluators are biased in favor of connected candidates, this is expected to encourage candidates with a connection in the committee to apply. Moreover, we would expect connected candidates to be negatively selected among applicants. On the other hand, if connections reduce information asymmetries, their impact would depend on the relative quality of candidates, particularly in dimensions that are observed more accurately by connected evaluators. Furthermore, we also analyze the impact of connections on researchers' chances of success. In what follows, given that we find that the impact of coauthors and colleagues is practically identical in empirical terms, we consider jointly both types of connections.[19]

[18] Information of affiliation is only available for evaluators based in Italy.

[19] Results disaggregated by coauthor and colleague are available upon request.

5.1 Applications

We estimate the causal impact of connections on researchers' application decisions. Researchers who have a connection in the evaluation committee tend to have a stronger research profile and, presumably, might also differ in some unobserved dimensions. We identify exogenous variations in the availability of a connection in the committee exploiting the random selection of committee members. We compare the application behavior of researchers who initially have similar chances of having a connection in the committee but, due to the random draw, differ in terms of the actual number of connections that they end up having in the evaluation committee:

$$y_{i,c} = \beta_0 + \beta_1 Connections_{i,c} + \mathbf{D_{i,c}}\boldsymbol{\beta_2} + \mu_c + \epsilon_{i,c}, \tag{6}$$

where $y_{i,c}$ is a dummy variable that takes value one if researcher i applies for a qualification in exam c (e.g.: qualification for an associate professorship in Econometrics). $\mathbf{D_{i,c}}$ represents a set of indicator variables for the number of connections that researcher i expects to have in committee c before the random selection takes place.[20] $Connections_{i,c}$ indicates the number of committee members selected in the initial random draw who have co-authored with the candidate or who are affiliated to the same institution (typically zero or one). A few evaluators (9%) resigned and were replaced by other (randomly chosen) eligible evaluators and, as result, the number of connections in the initial committee might differ slightly from the final composition of the committee at the time of the evaluation. Therefore, in the baseline specification coefficient β_1 captures the so-called intention-to-treat effect (ITT). In some additional specifications, we estimate the average treatment effect on the treated (ATET), instrumenting the final composition of committees using the initial composition that was determined by the random draw.

[20] We have computed the expected committee composition using one million simulated draws, taking into account the composition of the corresponding pools of eligible evaluators and the rules of the draw. We have then rounded it to two decimal places and created indicator variables for each value. All results, available upon request, are practically identical if we control for the expected number of connections using a linear specification instead of a set of dummies.

In order to increase the accuracy of the estimation, we include in the equation a set of exam fixed effects (μ_c), accounting for possible differences in the average success rate across different fields and positions. In some specifications, we also control for the set of predetermined individual characteristics and proxies of quality described in Table 2 ($\mathbf{X_i}$). In all regressions, standard errors are clustered at the field level, thus reflecting that evaluations within each field are done by the same committee.

The key identifying assumption of the analysis is that the composition decided by the initial random draw should not be correlated with any relevant observable or unobservable characteristic of researchers. The way in which the randomization was implemented suggests that there was little room for manipulation. Nonetheless, we explicitly test the randomness of the assignment. We estimate a specification similar to equation (6), but we now consider as dependent variables all observable predetermined characteristics of individual i (x_i). We perform this estimation on the sample of researchers who had pre-registered for the evaluation. As shown in Table 4, the results from these randomization tests are consistent with the assignment being random. Researchers who obtain, through luck of the draw, a connection in the evaluation committee are statistically similar to other researchers. While there are 10 coefficients that capture the correlation between the random shock to committee composition and researchers' characteristics, only one of these coefficients is statistically significant at the 10% level. The existence of random assignment is confirmed by the corresponding F-test for the joint significance of the estimates.

Table 5 reports the main estimates from equation (6). Researchers are significantly less likely to apply when they are assigned, through luck of the draw, to a committee that includes a connection. The presence of a co-author or a colleague in the initial committee decrease the probability of applying by 2.7 p.p., relative to an average application rate of 86% (column 1). These estimates become slightly more precise but they are statistically similar when we control for predetermined individual characteristics and observable productivity (column 2). The average treatment effect on the treated is slightly larger in absolute terms than the intention-to-treat estimate but the magnitudes are statistically similar (column 3).

We also analyze how application decisions vary depending on researchers' observable quality (columns 4-6). We split the sample in three groups based on researchers' publication record. In science, technology, engineering, mathematics and medicine (STEMM fields), we classify prospective applicants based on their total Article Influence Score and in social sciences and humanities we use the number of A-journal publications. The impact of connections on applications is driven by the decisions of researchers with weaker research profile. Connections do not have any significant impact on the application decisions of researchers in the top tercile but, for researchers in the lowest tercile, the presence of a co-author or a colleague in the committee decreases the likelihood to apply by about 6.2 p.p.

5.2 Bias vs. information

The presence of a co-author or a colleague in the committee decreases the probability that researchers will apply if they have a weak research profile. According to our theoretical framework, this pattern is consistent with two possible explanations. Some researchers perhaps withdraw their application because they anticipate that the connected evaluator might be negatively biased ($B_{i,connected} < 0$). This hypothesis is probably more plausible in the case of colleagues than in the case of coauthors. For instance, in some universities faculty members may be associated to different chairs that hold long-standing rivalries. Alternatively, the lower application rate of connected candidates may reflect a reduction in information asymmetries, either on the evaluator's side or on the candidate's side. Evaluators perhaps observe more accurately the quality of connected researchers ($\sigma_{i,connected} < \sigma_{i,unconnected}$) or, similarly, researchers may be better informed about the criteria of connected evaluators ($\gamma_{i,connected} < \gamma_{i,unconnected}$). Either way, the availability of more accurate information might discourage connected researchers with a weak research profile from applying.

Next we try to disentangle between the two alternative explanations, negative bias or lower information asymmetries. We use three sources of information: (i) the final outcome of the evaluation, (ii) the individual reports of committee members and (iii) the performance of researchers who withdrew their application in a subsequent round of the national scientific qualification.

5.2.1 Evaluations

To disentangle why connected researchers are more likely to withdraw their application, we compare the success rate of connected and unconnected researchers who received an evaluation. We estimate the following equation on the sample of researchers who did not withdraw the application:

$$y_{i,c} = \beta_0 + \beta_1 Connections_{i,c} + \mathbf{D_{i,c}}\boldsymbol{\beta_2} + \mathbf{X_i}\beta_5 + \mu_c + \epsilon_{i,c}, \tag{7}$$

where the dependent variable is an indicator that takes value one if the candidate qualifies and $\mathbf{X_i}$ includes all observable predetermined characteristics. This provides a consistent estimate of the causal impact of connections on assessments under the assumption that the set of controls fully accounts for any systematic differences in the quality of connected and unconnected candidates. However, given that committee composition affected application decisions, it is not possible to rule out selection based on unobservables characteristics. Connected researchers are positively selected in terms of their observable characteristics, and they might be also positively selected in terms of some relevant characteristics that are not observable to the econometrician. If this is the case, these estimates provide only an upper bound of the true impact of connections on evaluations.

Candidates with a connection in the committee are 6.6 p.p. (17%) more likely to qualify than other final candidates with comparable research outputs (Table 6, column 1). We observe a similar picture if we consider instead the total number of positive votes received by the candidate. The presence of a coauthor or a colleague in the committee increases the number of favorable votes by 0.32 (15%). The premium associated with connections does not vary depending on the research quality of candidates (columns 3-5).

5.2.2 Individual evaluation reports

Next we turn to the information provided by evaluators' individual assessments. First, we compare assessments made by different evaluators of the same candidate:

$$y_{i,j,c} = \beta_0 + \beta_1 Connection_{i,j} + \mu_i + \lambda_j + \epsilon_{i,j,c}, \tag{8}$$

where $y_{i,j,c}$ is a dummy variable that takes value one if evaluator j voted in favor of candidate i's application in qualification exam c. $Connection_{i,j}$ is a dummy variable indicating whether the candidate and the evaluator have coauthored in the past or they are based in the same institution. A set of application fixed effects (μ_i) controls for potential differences in the characteristics of candidates. In some specifications we also include evaluators' fixed effects (λ_j), which capture any potential differences in grading standards across evaluators. Coefficient β_1 captures the differences in the assessments received by each candidate from connected and unconnected evaluators, which might reflect the potential existence of differences in their evaluation criteria or in the available information.

Candidates are 3.9 p.p. (9%) more likely to get a positive vote from a colleague or a coauthor, relative to the assessments they receive from other committee members (Table 7, column 1). Controlling for the potential differences in the grading standards across evaluators does not affect these results (column 2). We also consider how this connection premium varies depending on the observable research output of candidates (columns 3-6). The premium is always positive, but it is slightly larger for candidates of lower quality.

The nature of the decision-making may actually have biased these estimates down. A high fraction of committees reach unanimous decisions, suggesting that there may be less disagreement reflected in these final verdicts than there would have been at interim stages. Nonetheless, given that these estimates are significantly positive, the evidence does not support the hypothesis that evaluators are negatively biased against their co-authors or their colleagues.

In addition to the direction of the vote, the content of individual reports may also provide information on the amount of information that was available to evaluators. In particular, the length of reports might reflect the extent to which the evaluator was initially informed and also the effort and the amount of time that she dedicated to collect additional information. We estimate equation (8) using as a left-hand side variable the length of individual reports, measured in words. As in the previous exercise, these estimates are likely to be a lower bound of the true gap, given that committee members share information and discuss evaluations before writing their final report.

While the average evaluation contains about 176 words, significantly longer assessments are provided by connected evaluators (Table 8, column 1). Co-authors and colleagues write on average 20 (12%) more words than other evaluators. This gap decreases to 12 words (7%) when we include evaluators' fixed effects, suggesting that an evaluator who has a connection to one particular candidate tends to write long assessments about all other candidates for the same position, but she writes even longer reports evaluations of their connections (column 2). As shown in columns 3-5, the gap does not vary significantly depending on the quality of candidates. We also explore whether the gap in length somehow reflects that evaluators tend to assess more positively connected candidates. This does not seem to be the case since evaluators write longer reports than other committee members independently of whether the individual assessment is positive or negative (columns 7 and 8). Overall, the gap in the length of reports is consistent with the hypothesis that connected evaluators are better informed about the quality of their co-authors and colleagues.

5.2.3 Future performance

The information from evaluators' assessments of candidates shows that connected candidates enjoy a premium. The existence of this connection premium is hardly surprising and is consistent with a number of previous empirical studies (Perotti 2002, Combes et al. 2008, Zinovyeva and Bagues 2015). But even if connected candidates benefit from connections, this does not necessarily imply that connected researchers who did not apply would have also enjoyed a positive premium. Perhaps the presence of a connection in the committee induced some researchers to withdraw their application precisely because they expected the connected evaluator to be negatively biased.

In order to examine the potential existence of a negative bias against connected researchers who withdrew their application, ideally we would like to have information on the assessments that they would have received had they applied. Unfortunately this information does not typically exist, and this case is no exception. Instead, we use the information provided by researchers' performance in the second round of the qualification exams, which took place the following year. In this second round, only those researchers who had not participated in the previous evaluation were allowed to apply. Most importantly, the composition of committees did not change between the first and the second round. Therefore, if the reason why connected researchers withdrew their application was that they anticipated some negative bias, this negative bias should also play a role in their decision to apply in the second round and in the assessments that they receive.

Relative to other researchers who also withdrew their application, connected researchers are more likely to apply in the second round and are also more likely to succeed (Table 9). Researchers with a coauthor or a colleague in the committee have a 4.1 p.p. (11%) higher probability of reapplying and are 9.4 p.p. (17%) more likely to succeed. This pattern is not consistent with the existence of a bias against connected researchers who withdrew their application in the first round. If anything, it suggests that the withdrawal was intended to improve the timing of the application.

In sum, the presence of a co-author or a colleague in an evaluation committee has a positive impact on the assessments that candidates receive. Connected candidates tend also to receive longer evaluation reports, suggesting that evaluators are better informed about their quality. Overall, the evidence suggests that connections are associated to a positive bias in assessments and also to lower information asymmetries. The latter effect seems to have a larger impact on application decisions, inducing some connected candidates with a weak research portfolio to postpone their application.

5.3 The overall impact of connections

The presence of connections in committees reduces the probability that researchers apply but it improves the assessments that they receive. Next we investigate the net impact of connections on candidates' chances of success, both in the short-term (first round of evaluations) and in the mid-term (second round of evaluations).

5.3.1 Short-term effect

We compare the success rate of connected and unconnected researchers in the first round of national qualification evaluations, exploiting the random assignment of evaluators to committees. We estimate equation (6) using as dependent variable an indicator which takes value one if pre-registered candidate i qualifies in examination c and value zero if the candidate failed or withdrew the application. As shown in column 1 of Table 10, the presence of a co-author or a colleague in the committee increases by 3.9 p.p. the probability of success of pre-registered candidates (or by 11% relative to the baseline success rate of 34%). The inclusion of individual controls increases threefold the explained variation in the dependent variable – the adjusted R-squared increases from 11% to 31% – but, as expected, it does not affect significantly the point estimates (column 2). The estimates are slightly larger, around 4.5 p.p., although statistically similar, when we instrument the final composition of the committee using the initial one (column 3). We also examine how the impact of connections on success varies depending on researchers' observable research productivity (columns 4-6). Good researchers benefit more from connections. Researchers in the top (bottom) tercile experience a 5.3 p.p. (3.0 p.p.) increase in their success rate when the committee includes a co-author or a colleague.

Connected candidates are significantly less likely to apply but they have significantly higher unconditional success rates. This necessarily implies that their chances of failing an exam, and therefore receiving a 2-year ban on re-applying, must be substantially lower. The probability that candidates with a co-author or a colleague in the committee apply and receive a negative assessment is 7.5 p.p. lower (column 7). Candidates with a weaker research profile benefit more from this decrease in failure rates. In the bottom tercile, the failure rate of connected candidates is 9.2 p.p. lower than the failure rate of other candidates, compared to a decrease of 6.1 p.p. for connected candidates in the top tercile (columns 4-6).

In sum, the extent to which candidates are affected by the presence of a connection in the committee depends on the quality of these same candidates. Top candidates face a larger increase in success rates. On the other hand, candidates with a relatively weaker research profile experience a larger decrease in application rates.

5.3.2 Mid-term effect

One of the advantages of not applying when failure is likely is the possibility of applying in the following round. To account for these longer-term effects, we analyze the impact of connections considering jointly the first and the second round of qualification exams. First, we examine the impact on applications. We estimate equation (6) using as left-hand variable an indicator that takes value one if candidate i applied either in the first or in the second round (Table 11, columns 1-4). On average, connections decrease application rates over the two rounds by 1.2 p.p. This is roughly one third of the impact on applications in the first round, indicating that the effect of connections on applications may be partially explained by connected candidates postponing their application for one year.

We also examine the overall impact of connections on success rates in both rounds (columns 5-8). The positive impact of connections is larger when we also take into account their impact in the second round. Considering both rounds, connected researchers are 6.2 p.p. more likely to qualify, compared to 4.5 p.p. in the first round.

Finally, we analyze the impact on failure rates (columns 9-12). The presence of a connection in the committee decreases the failure rate of connected candidates by 7.4 p.p. This effect is similar to the impact of connections on candidates' failure rate in the first round, and again it is larger for candidates with relatively lower research quality (9.0 p.p. vs 6.1 p.p.).

6 Conclusions

In this paper we study the role of connections in scientific committees exploiting the exceptional evidence provided by scientific evaluations in Italy. The impact of connections depends crucially on the research quality of researchers. Candidates with a strong research profile benefit from connections mostly by having higher chances of success. Researchers in the top tercile in terms of their research output are 5 p.p. more likely to succeed when the committee includes a coauthor or a colleague. Weaker researchers also benefit from connections by not making costly errors in application decisions. Researchers in the bottom tercile are 6 p.p. less likely to apply when the evaluation committee includes a co-author or a colleague and their chances of success are 3 p.p. higher. As a result, the probability that they fail the evaluation is 9 p.p. lower. Evidence from a subsequent round of evaluations suggests that, by postponing their application, weak researchers with a connection in the committee benefit also from higher success rates in the future. Overall, the evidence is consistent with the existence of a bias in favor of connected candidates and also with the notion that connections reduce information asymmetries.

Our analysis shows that, beyond their impact on evaluations, connections are also an important source of information for researchers application decisions. The analysis also provides strong evidence that self-selection might be an important source of concern for empirical studies that analyze evaluation biases. If prospective candidates can anticipate committee composition, this may affect their decision to apply. The direction of self-selection is difficult to predict and it will depend on the strength of evaluation biases, the degree of information asymmetries, and the quality of candidates. Selection might bias estimates if the econometrician can only observe the identity of actual candidates. This methodological problem is not limited to the analysis of connections in academia; it might be also relevant more generally in studies assessing evaluation biases in the labor market related to gender, ethnic group, or social ties (e.g. Fernandez and Weinberg 1997; Goldin and Rouse 2000; Petersen, Saporta and Seidel 2000, 2005). In such studies, ideally it would be convenient to consider not only actual applicants but also prospective ones.

Finally, our study also provides information that might be useful for the design of scientific evaluations. The system of national evaluations that was recently introduced in Italy is characterized by a large degree of transparency aimed at increasing meritocracy. However, publicizing CVs and evaluation reports is not sufficient to completely eliminate the connection premium. We still find that connected researchers are 4.5 p.p. (13%) more likely to qualify, although this figure is much lower than the connection premium observed in other countries where qualification exams are less transparent.[21] Moreover, the design of the system provides an additional advantage for connected candidates. Allowing candidates to withdraw their application after committee members have been selected helps connected candidates to take more informed application decisions and avoid costly failures.

[21] For instance, using data for Spain, Zinovyeva and Bagues (2015) find that the presence of a co-author or a colleague in the committee increases candidates' chances of qualifying by around 50%.

References

Azoulay, Pierre, Joshua Graff Zivin and Jialan Wang (2010), "Superstar Extinction", *Quarterly Journal of Economics*, 125(2), pp. 549-589.

Brogaard, Jonathan, Joseph Engelberg and Christopher A. Parsons (2014), "Network Position and Productivity: Evidence from Journal Editor Rotations," *Journal of Financial Economics*, 111 (1), pp: 251-270.

Combes, Pierre-Philippe, Laurent Linnemer and Michael Visser (2008), "Publish or peer-rich? The role of skills and networks in hiring economics professors," *Labour Economics*, Vol. 15, pp. 423-441.

Blau, Francine D., Janet M. Currie, Rachel T.A. Croson and Donna K. Ginther (2010), "Can Mentoring Help Female Assistant Professors? Interim Results from a Randomized Trial," *American Economic Review*, 100(2), 348-352.

Durante, Ruben, Giovanna Labartino, Roberto Perotti (2011), "Academic Dynasties: Decentralization and Familism in the Italian Academia", NBER WP 7572.

Fernandez, Roberto and Nancy Weinberg (1997), "Sifting and Sorting: Personal Contacts and Hiring in a Retail Bank," *American Sociological Review* 62(6) pp. 883-902.

Goldin, Claudia and Cecilia Rouse (2000), "Orchestrating impartiality: The impact of 'blind' auditions on female musicians," *American Economic Review* 90(4), 715-741.

Hamermesh, Daniel, and Peter Schmidt (2003), "The Determinants of Econometric Society Fellows Elections," Econometrica 71(1), 399-407.

Hastings, Justine, Christopher A. Neilson, and Seth D. Zimmerman (2015),"The Effects of Earnings Disclosure on College Enrollment Decisions," NBER Working Paper No. 21300.

Heckman, James, and Peter Siegelman (1993), "The Urban Institute Audit Studies: Their Methods and Findings," in *Clear and Convincing Evidence: Measurement* *of Discrimination in America*, ed. Fix and Struyk, 187-258. Washington, D.C.: The Urban Institute Press.

Hoxby, Caroline and Christopher Avery (2014), “The Missing ”One-Offs”: The Hidden Supply of High-Achieving, Low-Income Students,” Brookings Papers on Economic Activity, Spring, pp 1-50.

Li, Danielle (2011), “Information, Bias, and Efficiency in Expert Evaluation: Evidence from the NIH”, mimeo, MIT.

Laband, David N. and Michael J. Piette (1994), “Favoritism versus Search for Good Papers: Empirical Evidence Regarding the Behavior of Journal Editors”, *Journal of Political Economy*, Vol. 102(1), pp. 194-203.

Milkman, Katherine L., Modupe Akinola and Dolly Chugh (2015), “What Happens Before? A Field Experiment Exploring How Pay and Representation Differentially Shape Bias on the Pathway Into Organizations,” *Journal of Applied Psychology*, advance online publication, http://dx.doi.org/10.1037/apl0000022

OECD (2013), Education at a Glance 2013: OECD Indicators, OECD Publishing. http://dx.doi.org/10.1787/eag-2013-en.

Oettl, Alexander (2012), “Reconceptualizing Stars: Scientist Helpfulness and Peer Performance,” *Management Science*, Vol. 58(6), pp. 1122-1140.

Perotti, Roberto (2002), “The Italian University System: Rules vs. Incentives”, paper presented at the First Conference on Monitoring Italy, ISAE, Rome.

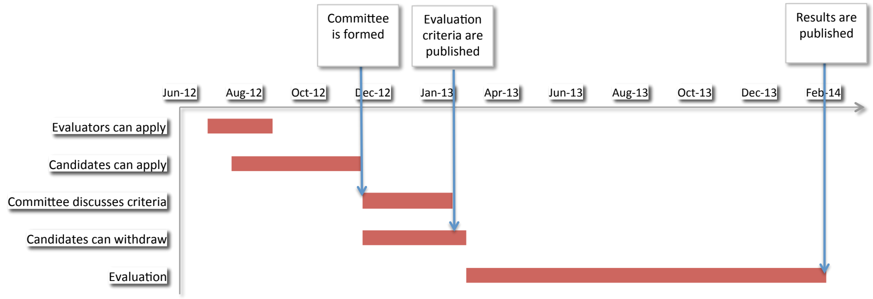
Petersen, Trond, Ishak Saporta and Marc-David Seidel (2000), “Offering a job: Meritocracy and social networks,” *American Journal of Sociology*, 106(3) pp. 763-816

Petersen, Trond, Ishak Saporta and Marc-David Seidel (2005) “Getting hired: Sex and race,” *Industrial Relations* 44(3), 416-443.

Sandström, Ulf and Martin Hällsten (2008), “Persistent nepotism in peer-review,” *Persistent nepotism in peer-review*, Vol. 74(2) pp. 175-189.

Zinovyeva, Natalia and Manuel Bagues (2015), "The Role of Connections in Academic Promotions", *American Economic Journal: Applied Economics*, 7(2): 264-92.

Figure 1: Timeline of the evaluation

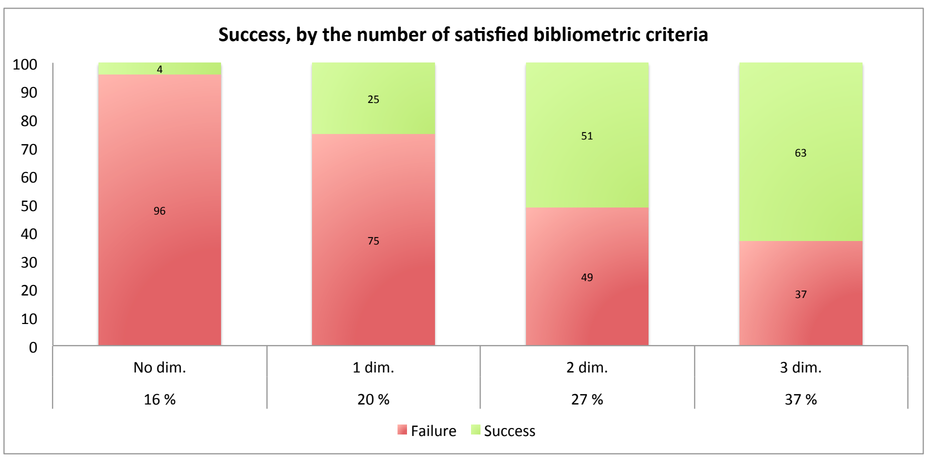


Note: The timeline is for Economics, discipline 13/A1.

Figure 2: Sample Individual Evaluation

DOE John
The candidate PINCO PALLO has been Ricercatore universitario at the Università di PISA since 2006. His scientific work is concerned with the development of democracy, including a monograph on the role of public opinion in political thought and a series of contributions concerning English and Anglo-American thought and developments from the 17th through 19th centuries, with special reference to Edmund Burke. The candidate is a member of the "Re-Imagining Democracy in the Mediterranean, 1750-1860" project, based at the University of Oxford. The candidate has a significant number of international conference participations, among which those in which the English have invited him to speak about Burke are perhaps the most indicative of a strong international reputation. In terms of specific contributions, the "silent guest" metaphor is particularly significant in explaining how Burke plays out in the history of Italian political thought. The candidate scores above the median on two of the three indicators of impact and has substantial relevant teaching experience. On the basis of the application submitted, the candidate merits approval of the request for the abilitazione scientifica.

Figure 3: Success rate and bibliometric measures



Note: Actual candidates have been classified in four groups, depending on the number of dimensions where their productivity is above the median in the corresponding category.

Table 1: Descriptive statistics – Eligible evaluators

	1	2	3	4
	Mean	Std. Dev.	Min	Max
Based in Italy (N=5,876):				
Female	0.20	0.40	0	1
All publications	131	104	4	957
- Articles	73	85	0	920
- Books	8	10	0	139
- Book chapters	22	26	0	455
- Conference proceedings	20	37	0	401
- Patents	0.42	2.44	0	88
- Other	7	23	0	675
Average Article Influence Score	1.18	0.73	0.1	9.65
A-journal articles	11	16	0	207
Based abroad (N=1,365):				
Female	0.12	0.32	0	1

Notes: Article Influence Score is defined for publications by professors in STEMM fields. A-journal articles are defined for publications by professors in the social sciences and humanities.

Table 2: Descriptive statistics – Applications

	1	2	3	4	5	6	7
			Position		Coauthor or colleague		
			FP	AP	Yes	No	
	Initial set of applications (N=69,020)						
	Mean	St.Dev.	Mean	Mean	Mean	Mean	p-value
Number of applications	1.49	1.08	1.52	1.48	1.34	1.53	0.000
Application order	0.5	0.29	0.5	0.5	0.46	0.51	0.000
Individual characteristics:							
Female	0.38	0.49	0.31	0.41	0.39	0.38	0.002
Age	44	8	49	43	0.05	-0.01	0.000
Permanent university position:	0.55	0.5	0.74	0.47	0.74	0.52	0.000
- same field	0.75	0.43	0.77	0.74	0.79	0.74	0.000
Quality indicators:							
CV length (pages)	16	67	20	14	0.08	-0.02	0.000
All Publications:	64	67	89	53	0.08	-0.02	0.000
- Articles	37	51	53	30	0.07	-0.01	0.000
- Books	2	5	3	2	0.01	-0.00	0.509
- Book chapters	7	12	10	6	0.06	-0.01	0.000
- Conference proceedings	10	20	14	8	0.07	-0.01	0.000
- Patents	0.24	1.65	0.35	0.19	0.00	-0.00	0.936
- Other	7	22	8	7	-0.02	0.00	0.004
Average number of coauthors	6	18	6	6	0.01	-0.00	0.229
First-authored	0.22	0.2	0.22	0.22	-0.02	0.00	0.069
Last-authored	0.12	0.16	0.15	0.11	0.03	-0.01	0.002
Average Article Influence Score	1.31	0.97	1.31	1.30	-0.01	0.00	0.296
A-journal articles	4	7	6	3	0.09	-0.01	0.000
	Final set of applications (N=59,150)						
Production in the previous 10 years:							
Social Sciences and Humanities:							
- Articles	20	17	25	18	0.16	-0.02	0.000
- A-journal articles	3	4	3	2	0.09	-0.01	0.000
- Books	2	3	3	2	0.02	-0.00	0.367
Sciences:							
- Articles	37	45	46	32	0.06	-0.01	0.000
- Citations	60	102	77	52	0.05	-0.01	0.000
- H-index	11	7	13	10	0.09	-0.02	0.000
Above the median in 3 indicators	0.38	0.48	0.42	0.36	0.46	0.36	0.000
Below the median in 3 indicators	0.16	0.36	0.13	0.17	0.12	0.17	0.000

Notes: Article Influence Score is defined for publications by professors in STEMM fields. A-journal articles are defined for publications by professors in the social sciences and humanities. Columns 5-6 provide information for the subset of applicants who had a connection in the committee and the subset who did not. Column 7 reports the p-value for the t-test of difference in means between the two groups. In columns 5-6 productivity indicators and age are normalized at the exam level.

Table 3: Descriptive statistics – Outcomes

	1	2	3	4	5	6	7	8
				Position		Coauthor or colleague		
				FP	AP	Yes	No	
	N	Mean	St.Dev.	Mean	Mean	Mean	Mean	p-value
Initial set of applications in the 1^{st} *wave*								
Withdraw	69,020	0.14	0.35	0.16	0.13	0.17	0.14	0.000
Fail	69,020	0.49	0.50	0.48	0.50	0.34	0.52	0.000
Qualify	69,020	0.37	0.48	0.36	0.37	0.49	0.34	0.000
Final set of applications in the 1^{st} *wave*								
Qualify	59,150	0.43	0.49	0.43	0.43	0.59	0.40	0.000
Unanimous decision	58,948	0.86	0.35	0.84	0.86	0.86	0.86	0.813
Individual evaluations								
Length of individual evaluations	294,656	176	277	203	164	193	175	0.000
Positive votes	294,656	0.45	.50	0.46	0.44	0.64	0.44	0.000
Set of withdrawn applications in the 1^{st} *wave*								
Reapply in the 2^{nd} wave	9,870	0.37	0.48	0.32	0.40	0.44	0.36	0.000
Set of resubmitted applications in the 2^{nd} *wave*								
Qualify in the 2^{nd} wave	3,647	0.58	0.49	0.59	0.57	0.67	0.55	0.000

Notes: We observe 99.7% of individual evaluations (294,656 out of 295,666 evaluations).

Table 4: Randomization test

	1	2	3	4	5	6	7	8	9	10
Dependent variable:	Female	Age	Perm.pos. same field	Perm.pos., other field	Appl. order	CV length	Publ.	A-journal articles	Total AIS	Coauthors
Connection in committee	0.005	0.026*	0.002	0.002	-0.000	-0.025	0.001	-0.005	-0.011	-0.015
	(0.006)	(0.014)	(0.008)	(0.005)	(0.004)	(0.017)	(0.019)	(0.011)	(0.018)	(0.019)
Observations	69,020	69,020	69,020	69,020	69,020	69,020	69,020	69,020	69,020	69,020

Notes: OLS estimates. All regressions include exam fixed effects and set of dummy variables for the expected number of connections in the committee (192 dummies). Dependent variables in columns 2, 5-10 are normalized at the exam level.
Standard errors are clustered at the committee level. *** denotes significance at 1%, ** significance at 5% and * significance at 1%.

Table 5: Application decisions

	1	2	3	4	5	6
				Research productivity:		
Sample:	All	All	All	High	Medium	Low
	ITT	ITT	ATET	ATET	ATET	ATET
Connection in committee	-0.027***	-0.027***	-0.030***	-0.009	-0.019**	-0.062***
	(0.005)	(0.005)	(0.005)	(0.006)	(0.008)	(0.011)
Individual controls	No	Yes	Yes	Yes	Yes	Yes
Observations	69,020	69,020	69,020	21,443	21,800	25,777
Adjusted R-squared	0.045	0.118	0.119	0.146	0.120	0.138
Mean, no connections	0.862	0.862	0.862	0.935	0.869	0.799
Connection effect, %	-3.110	-3.155	-3.457	-0.920	-2.155	-7.760

Note: The dependent variable takes value one whenever the researcher keeps the application for the evaluation. Columns 1 and 2 report results from an OLS estimation where the main right-hand side variables reflect the initial composition of the committee, providing the intention-to-treat effects (ITT). Columns 3-6 report results from instrumental variables estimations where the final composition of the committee has been instrumented using the outcome of the initial random draw, providing the average treatment effect on the treated (ATET).
Research productivity is measured by the total Article Influence Score in STEMM fields and by publications in A-journals in the social sciences and humanities.
All regressions include exam fixed effects and a set of dummy variables for the expected number of connections in committee. Individual controls include position, university, and all variables in the upper panel of Table 2.
Standard errors are clustered at the committee level. *** denotes significance at 1%, ** significance at 5% and * significance at 1%.

Table 6: Evaluations

	1	2	3	4	5
Dependent variable:	Qualify	Positive votes	Qualify	Qualify	Qualify
Sample:	All final candidates		Research productivity:		
			High	Medium	Low
Connection in committee	0.066***	0.319***	0.066***	0.062***	0.064***
	(0.006)	(0.028)	(0.009)	(0.008)	(0.010)
Individual controls	Yes	Yes	Yes	Yes	Yes
Observations	59,150	59,150	20,028	18,855	20,267
Adjusted R-squared	0.422	0.451	0.380	0.373	0.381
Mean, no connections	0.399	2.084	0.586	0.446	0.186
Connection effect, %	16.6	15.3	11.2	13.9	34.1

Notes: OLS estimates. All regressions include exam fixed effects and a set of dummy variables for the expected number of connections in committee. Individual controls include position, university, and all variables in the upper and lower panels of Table 2.
Research productivity is measured by the total Article Influence Score in STEMM fields and by publications in A-journals in the social sciences and humanities.
Standard errors are clustered at the committee level. *** denotes significance at 1%, ** significance at 5% and * significance at 1%.

Table 7: Evaluators' individual voting

	1	2	3	4	5
Sample:	All final candidates		Research productivity:		
			High	Medium	Low
Connection	0.039***	0.039***	0.030***	0.043***	0.047***
	(0.005)	(0.005)	(0.005)	(0.006)	(0.008)
Candidate fixed-effects	Yes	Yes	Yes	Yes	Yes
Evaluator fixed-effects	No	Yes	Yes	Yes	Yes
Observations	294,656	294,656	99,747	93,969	100,940
Number of applications	58,948	58,948	19,957	18,799	20,192
Adjusted R-squared	0.002	0.066	0.078	0.081	0.084
Mean, no connections	0.440	0.440	0.624	0.488	0.217
Connection effect, %	9.0	8.9	4.8	8.8	21.5

Notes: OLS estimates. Each observation represents evaluator j assessment of candidate i. The dependent variable is a dummy that takes value one if the evaluator votes in favor of the candidate.
Research productivity is measured by the total Article Influence Score in STEMM fields and by publications in A-journals in the social sciences and humanities.
Standard errors are clustered at the committee level. *** denotes significance at 1%, ** significance at 5% and * significance at 1%.

Table 8: Length of individual reports

	1	2	3	4	5	6	7
Sample:	All final candidates		Research productivity:			Received a positive vote	Received a negative vote
			High	Medium	Low		
Connection	20.2***	12.0***	12.0***	12.4***	9.2***	13.2***	5.6***
	(3.7)	(1.9)	(1.7)	(3.1)	(2.0)	(1.7)	(1.0)
Candidate fixed-effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Evaluator fixed-effects	No	Yes	Yes	Yes	Yes	Yes	Yes
Observations	294,656	294,656	99,747	93,969	100,940	131,613	163,043
Number of applications	58,948	58,948	19,957	18,799	20,192	30,722	36,638
Adjusted R-squared	0.001	0.347	0.415	0.306	0.381	0.421	0.343
Mean, no connections	175.5	175.5	182.8	178.5	165.6	187.6	166.0
Connection effect, %	11.5	6.9	6.3	7.0	5.6	7.1	3.4

Notes: OLS estimates. Each observation represents evaluator j assessment of candidate i. The dependent variable is the length of individual evaluation reports measured in words.
Research productivity is measured by the total Article Influence Score in STEMM fields and by publications in A-journals in the social sciences and humanities.
Standard errors are clustered at the committee level. *** denotes significance at 1%, ** significance at 5% and * significance at 1%.

Table 9: The impact of connections - 2^{nd} wave

	1	2
Dependent variable:	Reapply in the 2^{nd} wave	Qualify in the 2^{nd} wave
Sample:	Withdrawn applications in the 1^{st} wave	Resubmitted applications in the 2^{nd} wave
Connection in committee	0.041***	0.094***
	(0.014)	(0.025)
Observations	9,870	3,647
Adjusted R-squared	0.158	0.204
Mean, no connections	0.357	0.551
Connection effect, %	11.4	17.0

Notes: OLS estimates. All regressions include exam fixed effects and a set of dummy variables for the expected number of connections in committee. Individual controls include position, university, and all variables in the upper panel of Table 2.
Standard errors are clustered at the committee level. *** denotes significance at 1%, ** significance at 5% and * significance at 1%.

Table 10: Success and Failure

	1	2	3	4	5	6	7	8	9	10
Dependent variable:	Qualify						Fail			
				Research productivity:				Research productivity:		
Sample:	All	All	All	High	Medium	Low	All	High	Medium	Low
	ITT	ITT	ATET	ATET	ATET	ATET	ATET	ATET	ATET	ATET
Connection in committee	0.039***	0.041***	0.045***	0.053***	0.047***	0.030***	-0.075***	-0.061***	-0.066***	-0.092***
	(0.007)	(0.006)	(0.006)	(0.010)	(0.009)	(0.009)	(0.007)	(0.010)	(0.009)	(0.012)
Individual controls	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	69,020	69,020	69,020	21,443	21,800	25,777	69,020	21,443	21,800	25,777
Adjusted R-squared	0.111	0.307	0.307	0.336	0.274	0.255	0.237	0.295	0.220	0.205
Mean, no connections	0.344	0.344	0.344	0.548	0.387	0.149	0.518	0.387	0.482	0.650
Connection effect, %	11.3	12.0	13.2	9.7	12.1	19.9	-14.5	-15.9	-13.6	-14.1

Notes: In columns 1-6 the dependent variable takes value one if the candidate qualified and it takes value zero if the candidate failed or withdrew the application. In columns 7-10 the dependent variable is one if the candidate failed, and zero otherwise. Columns 1-2 report results from an OLS estimation where the main right-hand side variables reflect the initial composition of the committee, providing the intention-to-treat effects (ITT). Columns 3-10 report results from instrumental variables estimations where the final composition of the committee has been instrumented using the outcome of the initial random draw, providing the average treatment effect on the treated (ATET).
Research productivity is measured by the total Article Influence Score in STEMM fields and by publications in A-journals in the social sciences and humanities.
All regressions include exam fixed effects and a set of dummy variables for the expected number of connections in committee. Individual controls include position, university, and all variables in the upper panel of Table 2.
Standard errors are clustered at the committee level. *** denotes significance at 1%, ** significance at 5% and * significance at 1%.

Table 11: The effect of connections on two-period outcomes

	1	2	3	4	5	6	7	8	9	10	11	12
Dependent variable:	Apply in the 1^{st} or the 2^{nd} wave				Qualify in the 1^{st} or the 2^{nd} wave				Fail in the 1^{st} or the 2^{nd} wave			
Sample:	All	Research productivity:			All	Research productivity:			All	Research productivity:		
		High	Medium	Low		High	Medium	Low		High	Medium	Low
Connection in committee	-0.012***	-0.001	0.001	-0.036***	0.062***	0.058***	0.066***	0.053***	-0.074***	-0.061***	-0.066***	-0.090***
	(0.004)	(0.005)	(0.006)	(0.009)	(0.006)	(0.010)	(0.009)	(0.010)	(0.007)	(0.010)	(0.009)	(0.011)
Observations	69,020	21,443	21,800	25,777	68,453	21,272	22,651	25,530	68,453	21,272	21,651	25,530
Adjusted R-squared	0.088	0.117	0.090	0.111	0.312	0.338	0.284	0.279	0.240	0.299	0.225	0.193
Expected connections	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Individual controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Committee FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean, no connections	0.911	0.961	0.925	0.861	0.372	0.568	0.423	0.178	0.539	0.394	0.504	0.683
Connection effect, %	-1.3	-0.1	0.1	-4.1	16.6	10.3	15.7	30.0	-13.7	-15.4	-13.0	-13.1

Notes: The table reports results from instrumental variables estimations where the final composition of the committee has been instrumented using the outcome of the initial random draw, providing the average treatment effect on the treated (ATET).
Research productivity is measured by the total Article Influence Score in STEMM fields and by publications in A-journals in the social sciences and humanities.
All regressions include exam fixed effects and a set of dummy variables for the expected number of connections in committee. Individual controls include position, university, and all variables in the upper panel of Table 2.
Standard errors are clustered at the committee level. *** denotes significance at 1%, ** significance at 5% and * significance at 1%.