Does the Gender Composition of Scientific Committees Matter?

By Manuel Bagues, Mauro Sylos-Labini, and Natalia Zinovyeva*

ONLINE APPENDIX

Appendix A. Summary of the literature

In this section we present a brief summary of the literature that analyzes how the gender composition of academic committees affects the relative success rate of female candidates. Table A1 summarizes the existent studies along a number of key dimensions: the type of analyzed evaluation, the field in which the analyzed evaluation took place, the empirical method used to identify the causal impact of the committee composition on the evaluation outcome, the size of the sample as measured by the number of applications evaluated by committees, and the main result of each study.

* Bagues: Aalto University, Department of Economics, P.O. Box 21240, FI-00076 AALTO, Helsinki, Finland, manuel.bagues@aalto.fi. Sylos-Labini: University of Pisa, Department of Political Science, Via Serafini 3, Pisa, Italy, mauro.syloslabini@unipi.it. Zinovyeva: Aalto University, Department of Economics, P.O. Box 21240, FI-00076 AALTO, Helsinki, Finland, natalia.zinovyeva@aalto.fi.

TABLE A1—**Summary of the literature**

Paper	Type of evaluation	Field	Empirical method	Applications	Results
Broder (1993)	Grant applications	Economics	Application fixed effects	1,479	Opposite-sex preference
Steinpreis, Anders and Ritzke (1999)	Job applicants and tenure candidates	Psychology	Randomized field experiment	238	No significant difference
Jayasinghe, Marsh and Bond (2003)	Grant applications	Several	Application fixed effects	2331	No significant difference
Ellemers et al. (2004)	Work commitment of students	Several	Identification based on observables	212	Opposite-sex preference
Milkman, Akinola and Chugh (2015)	Prospective students request a meeting	Several	Randomized field experiment	6,548	No significant difference
Moss-Racusin et al. (2012)	Laboratory manager position	Life Sciences	Randomized field experiment	127	No significant difference
Abrevaya and Hamermesh (2012)	Paper submitted for publication	Economics	Application fixed effects	2,940	No significant difference
Casadevall and Handelsman (2014)	Selection of conference speakers	Microbiology	Identification based on observables	1,845	Same-sex preference
De Paola and Scoppa (2015)	Job applicants	Economics and Chemistry	Identification based on observables	2,279	Same-sex preference
Williams and Ceci (2015)	Job applicants	Several	Randomized field experiment	873	No significant difference

Note: $^{(*)}$ We classify the empirical strategy followed by De Paola and Scoppa (2015) as identification based on observables and not as randomized natural experiment due to the nature of the empirical strategy implemented by the authors. This paper studies promotions in the Italian university system that was in place between 2008 and 2011. In this system, four out of five members of the evaluation committee were randomly selected from a pool of eligible evaluators. In their analysis, the authors study the relationship between the success rate of male and female candidates and the gender composition of committees, unconditional on the gender composition of the pool of eligible evaluators. The consistency of the estimation relies on the implicit assumption that the relative quality of male and female applicants is unrelated to the degree of feminization of the pool of eligible evaluators.

Appendix B. Institutional background

There are several important differences between the Spanish and the Italian systems of centralized national evaluations. To facilitate the comparison, Table B1 summarizes the main features of the two systems.

TABLE B1—**Main features of the evaluation systems in Italy and Spain**

	Italy, Abilitazione Scientifica Nazionale, 2012-2014	**Spain, Habilitación, 2002-2006**
Eligibility requirement for candidates	None	None
Size of evaluation committees	5 evaluators	7 evaluators
Assignment to committees	Based on a random draw	Based on a random draw
Composition of committees	4 full professors based in Italian universities, 1 professor based abroad	In full professor exams, 7 full professors based in Spanish universities or public research centers. In associate professor exams, 3 full professors and 4 associate professors.
Constraints on randomization	No university can have more than one evaluator within a single committee.	Only one non-university researcher is allowed to be selected as a member of the committee for a given exam. Similarly, only one emeritus professor is allowed to be selected as a member of a given committee.
Minimum research quality requirement for evaluators	In STEMM disciplines, eligible professors should be above the median in their category and field in at least two of the following dimensions: (i) the number of articles published in scientific journals, (ii) the number of citations, (iii) and the H-index. In SSH disciplines, they should be above the median in at least one of the following dimensions: (i) the number of articles published in high impact scientific journals (so-called A-journals), (ii) the overall number of articles published in any scientific journals and book chapters, and (iii) the number of published books.	Eligible associate professors should have one *sexenio* and eligible full professors should have two *sexenios*. *Sexenios* are granted by the Spanish education authority on the basis of applicants' research output in any uninterrupted period of a maximum of six years.
Inclusion in the pool of eligible evaluators	Voluntary	Compulsory
Substitution of resigned evaluators	Based on a random draw	Based on a random draw
Voting rule	Qualified majority of 4	Simple majority
Number of qualifications granted by the committee	Unlimited	Limited by the number of available positions at the university level
Validity of a positive qualification	4 years (later extended to 6 years)	Unlimited
Penalization for a negative evaluation	2 years application ban	None
Application withdrawal	Up until two weeks after the evaluation criteria are publicized	Candidates can drop out from the process at any time
Evaluation	Evaluations are based solely on the material provided in candidates' application packages, consisting of CVs and selected publications.	Oral exams to full professor positions have two qualifying stages. In the first stage, candidates present their CVs. In the second stage, candidates present a piece of their research work. Exams to associate professor, in addition to these two stages, have an intermediate stage where candidates give a lecture on a topic randomly chosen from a syllabus proposed by the candidate.
Degree of transparency	The lists of potential and actual evaluators and candidates, as well as the lists of qualified candidates, are published online. Furthermore, the CVs of all participants and individual evaluation reports are published online. The evaluation agency also collects and publicizes information on the bibliometric indicators of candidates.	The lists of potential and actual evaluators and candidates, as well as the lists of qualified candidates, are published online.

Appendix C. Data

The data on the participants in Italian evaluations, including the CV of all eligible evaluators and all candidates, was available at the website of the Italian Ministry of Higher Education and Research. We extracted all the individual characteristics that we use in the analysis from these CVs. Information on tenured researchers' affiliation and the length of tenure was obtained from the Consortium of Italian universities (CINECA). Affiliation of nontenured researchers is from the most recent publication of the CV.

We also downloaded from the website of the Italian Ministry approximately 295,000 individual evaluation reports, five per each candidate. Due to the data collection problem, we are missing information on individual evaluations for 202 candidates. We are also missing 84 individual evaluation reports in three committees where evaluators abstained whenever there was a conflict of interest. We conducted a text analysis of the available individual evaluation reports and we identified approximately 9,000 different sentences that indicate the evaluator's decision to fail or to pass a given candidate.

The data on the participants in Spanish evaluations was collected from different sources, including the Spanish Ministry of Research and Science, Thomson Reuters (ISI) Web of Knowledge, the database of publications in Spanish language Dialnet, the European Patent Office and TESEO database on doctoral dissertations.[29]

Publications indexed in above sources are matched to the list of professors in Spain based on individuals' names and field of research. This process suffers from an important problem with homonymity since there are lots of common surnames in Spain. In addition to this, bibliographic databases often incompletely record authors' names (this especially concerns the data on publications before 2010 in the Web of Knowledge). Facing the choice between minimizing the number of false positives or the number of false negatives, we generally preferred the former. This means that, on the one hand, the individuals are authors of the outcomes assigned. On the other hand, we are unable to assign research outputs that have an incomplete record of authors' names.

Below we describe in detail the process of data collection in the case of Spain.

Spanish Ministry of Research and Science

The Spanish system of centralized examinations known as 'habilitación' was in place between 2002 and 2006. In total, 1,016 exams took place, around five per discipline. We restrict the sample in several ways. We exclude exams where the number of available positions was larger or equal than the number of candidates (two exams, both in Basque Philology) and disciplines where the number of potential evaluators was not large enough to form a committee (55 exams).[30] The final database includes 967 exams.

Information on candidates' and evaluators' first name, last name, tenure and ID number was retrieved from the website of the Ministry of Research and Science in July 2009 (`http://micinn.es`). Information on first names allows us to identify gender. In a few cases where it was not possible to assign gender based on first name, we searched online for a personal picture or document that would make it possible to assign gender.

The actual age of individuals is not observable. Instead, we exploit the fact that Spanish ID numbers contain information on their issue date to construct a proxy for the age of native individuals on the basis of his/her national ID number. In Spain, police stations are given a range of ID numbers which are assigned to individuals in a sequential manner. Since it is

[29] We would like to thank Stéphane Maraut and Catalina Martinez for kindly sharing the data on academic inventors who have patented their inventions in the European Patent Office. For a description of how the patent data was collected and matched to professors, see Maraut and Martínez (2014).

[30] In these cases, unfilled seats in the committee were filled with professors from related disciplines.

compulsory for all Spaniards to have an ID number by age 14, two Spaniards with similar ID numbers are likely to be of the same age (and geographical origin).[31] In order to perform the assignment, we first use registry information on the date of birth and ID numbers of 1.8 million individuals in order to create a correspondence table which assigns year of birth to the first four digits of ID number (ranges of 10,000 numbers). To test the precision of this correspondence, we apply it to a publicly available list of 3,000 court clerks, which contains both the ID number and the date of birth. In 95 percent of the cases the assigned age is within a three-year interval of the actual age. In order to minimize potential errors, whenever our age proxy indicated that a candidates for an associate professorship is less than 27 years old and a candidate for full professorship is less than 35 years old, we assign age a missing value. This proxy is also not defined for non-Spaniards (less than 1 percent of the sample). We imputed the missing age with the average age of individuals at the same discipline and rank (around 5 percent of the sample).

In 2006 the system of *habilitación* was replaced by a system known as *acreditación*, which is still in place. Under the *acreditación* system applicants aspiring for promotion are also required to be approved by a national review committee. These committees evaluate candidacies on a monthly basis and their decisions are published in the Official State Bulletin. We collected information on the identity of all candidates that qualified for a FP position before September 2013.

The Ministry provides information on affiliation and on tenure in the position for eligible evaluators. Given that most candidates to full professor positions are eligible evaluators themselves in exams to associate professor positions, it is possible to obtain their affiliation by matching the list of eligible evaluators with the list of candidates. Using this procedure, we were able to obtain the information on affiliation for 93 percent of candidates to full professor positions. We obtained the information on affiliation for the remaining 7 percent of candidates from the State Official Bulletin or directly from professors' CVs that can be found online.

ISI Web of Knowledge

We also collected information on the research output of eligible evaluators and candidates from the ISI Web of Knowledge.[32]

Information on scientific publications comes from the Thomson Reuters ISI Web of Knowledge (WoK). We consider publications published since 1972 by authors based in Spain, as well as the number of citations received by these publications before July 2009. The WoK database includes over 10,000 high-impact journals in the categories of Science, Engineering, Medicine and Social Sciences, as well as international proceedings coverage for over 110,000 conferences. For the purpose of this analysis, we considered all articles, reviews, notes and proceedings.

The assignment of articles to professors is not trivial. For each publication and author, WoS provides information on his/her surname and on his/her initial. In Spain, some surnames are very common (e.g., Garcia, Fernandez, Gonzalez), and this may create problems with homonymity. Moreover, unlike most other countries, individuals are assigned two surnames (paternal and maternal) and sometimes also several first names. When Spanish authors sign a paper they may do it with only their paternal or with their maternal surname, or they may hyphenate the two surnames. Authors may also sign using their first name, their middle

[31] There are a number of exceptions. For instance, this methodology will fail to identify the age of individuals who obtained their nationality when they were older than 14. Nevertheless, immigration was a rare phenomenon in Spain until the late 1990s. Additionally, some parents may have their children obtain an ID number before they are 14. This may be the case particularly after Spain entered in the mid 90s the Schengen zone and IDs became a valid documentation to travel to a number of European countries.

[32] We are grateful to the *Fundación Española para la Ciencia y la Tecnología* for providing us with access to the data.

name, or both.

We use the following matching procedure in order to deal with the above problems. First, we assign all publications and all professors in our sample to a broad disciplinary category. In order to attribute comparable disciplinary categories for publications and individuals, we aggregate disciplines defined by the Spanish Ministry and ISI disciplinary areas into the following categories: Agriculture; Chemistry; Biology; Geology; Physics; Mathematics and Computer Science; Engineering; Medicine, Veterinary and Pharmacology; Economics and Management; Psychology, Sociology and Political Science.[33] Second, in each broad disciplinary category we match publications with individuals in our database using the information on their surnames and initials.

Specifically, the publication is assigned to a professor in the list of eligible evaluators if it is in the same disciplinary category as the professor, and the author's surname and initial, as reported by ISI, coincide (i) with the first surname and the first name's initial of the professor, (ii) with the last surname and the first initial, (iii) with the first surname hyphenated with the second surname and the first initial. We also repeat stages (i) through (iii) substituting the first initial with the middle-name initial. If a given publication can be assigned to more than one possible match, the value of this publication is divided by the number of such possible matches.

Given that the propensity to publish differs substantially across the disciplines, we normalize the number of individual's publications to have zero mean and unit standard deviation among applicants to the same exam and among eligible evaluators of a given category in a given exam. The number of citations of each publication depends on the time elapsed between the publication date and the date when the number of received citations is observed. Therefore, we first normalize the number of citations that each publication receives by subtracting the average number of citations received by Spanish-authored articles published in the corresponding ISI disciplinary area in the same year and then dividing by the corresponding standard deviation. Next, for each individual in our database we calculate the average number of citations per publication. For individuals who have no ISI publications, this variable takes the minimum value in the corresponding discipline. Finally, similarly to the number of publications, we normalize the number of individual's citations per publication to have zero mean and unit standard deviation among applicants to the same exam and among eligible evaluators of a given category in a given exam.

Dialnet

Dialnet (`http://dialnet.unirioja.es`) is an open access bibliographic index created by the University of Rioja. It contains information on more than 8,000 journals and more than 3,5 million documents in Hispanic languages, including articles published in scientific journals, collective works and books. The database mainly covers publications in social sciences and humanities. Dialnet provides (in most cases) systematized information on individual authors' first name, paternal surname, maternal surname and affiliation, thus limiting potential concerns about homonymity.

We collected information on publications in Dialnet. Due to its lack of representativeness, we did not considered publications in Science and Engineering. We also excluded publications that appear in ISI Web of Science. We also restricted the set of journals considered to those which satisfy certain minimum research quality requirements (categories A, B or C) as established by the Integrated Scientific Journals Classification (CIRC) (Torres-Salinas et al., 2010). Similarly, we considered only books and collective volumes that are published by publishers

[33] In practice, apart from the case of journals *Science* and *Nature*, the ISI scientific categories are assigned to journals, not publications. In very rare cases a publication happened to be assigned to more than one broad disciplinary group.

that satisfy a minimum quality requirement. In particular, we used the EPUC-CSIC publisher list, which summarizes the names of the main publishers in social sciences and humanities in Spain and abroad (Giménez-Toledo, Tejada-Artigas and Mañana-Rodríguez, 2012). Publications that have been excluded from our study are mainly publications in working paper series, nonrefereed journals and volumes published by local universities (around 30 percent).

Teseo database on doctoral dissertations

Since 1977, PhD candidates in Spanish universities have registered their dissertation in the database TESEO, which is run by the Ministry of Education. We retrieved all the information available in this database from the website `https://www.educacion.gob.es/teseo` in May 2011. While registration is compulsory, according to Fuentes and Arguimbau (2010) TESEO includes information on approximately 90 percent of all dissertations read in Spain during this period. We observe information on 151,483 dissertations. TESEO provides the identity and affiliation of dissertations' authors, advisors and committee members. Approximately 40 percent of dissertations are female authored. Female supervisors are scarce and represent only 18 percent of the total. While 58 percent of the students they supervise are female, in the case of male advisors, 61 percent of their students are male.

We match TESEO data with the list of candidates and evaluators. In exams to full professor positions we are able to find the dissertation of 71 percent of candidates and 41 percent of evaluators. In exams to associate professor positions we observe the dissertation of 83 percent of candidates and 70 percent of evaluators. Missing information may be due to the fact that individuals (i) did their PhD abroad, (ii) defended their dissertation before 1977, (iii) there are spelling mistakes, (iv) the dissertation was not included in TESEO for unknown reasons (approximately 10 percent of all dissertations), or (v) there was a problem with homonymity (in our dataset 0.1 percent of individuals share the same name, middle name, paternal surname and maternal surname).

Each thesis has been classified by its author using the Unesco International Standard Nomenclature for Fields of Science and Technology. This system developed by Unesco includes more than two thousand six-digits categories.[34] 80 percent of dissertations provide this information. Approximately half of the authors select one six-digit category, 35 percent select two categories, and 15 percent select three or more categories. There are on average around one hundred dissertations per category. We use this information to construct a measure of individuals' research interests. In particular, we take into account every dissertation where an individual appears as an advisor, committee member or author. We were able to obtain information on the research interests of 98 percent candidates to full professor positions, 94 percent of candidates to associate professor positions, 98 percent of eligible full professors and 96 percent of eligible associate professors.

[34] Available at `http://unesdoc.unesco.org/images/0008/000829/082946eb.pdf`

TABLE C1—**Descriptive statistics – Eligible evaluators**

	1	2	3	4	5	6	7	8	9	10	11	12
	Italy				**Spain**							
	Full professors				Full professors				Associate professors			
	All	Male	Female	p-value	All	Male	Female	p-value	All	Male	Female	p-value
Female	0.20				0.14				0.35			
Tenure in position	13	0.07	-0.27	0.000	13	0.05	-0.33	0.000	10	0.01	-0.03	0.000
Age	-				52	0.00	-0.03	0.010	45	0.01	-0.02	0.001
All Publications:	131	0.04	-0.17	0.000	34	0.02	-0.14	0.000	14	0.05	-0.09	0.000
- Articles	73	0.05	-0.20	0.000	30	0.02	-0.14	0.000	12	0.05	-0.10	0.000
- Books	8	0.05	-0.19	0.000	1	0.01	-0.06	0.000	0.44	0.01	-0.02	0.000
- Book chapters	22	0.02	-0.07	0.003	3	0.01	-0.07	0.000	1	0.01	-0.01	0.003
- Conference proceedings	20	-0.00	0.02	0.552	-				-			
- Patents	0.42	0.00	-0.01	0.522	0.10	0.00	-0.02	0.001	0.04	0.01	-0.02	0.000
- Other	7	-0.00	0.00	0.909	-				-			
Total Article Influence Score	133	0.04	-0.24	0.000	32	0.01	-0.09	0.000	12	0.03	-0.07	0.000
A-journal articles	11	0.05	-0.14	0.000	4	0.03	-0.18	0.000	2	0.04	-0.06	0.000
All Publications, previous 10 years	74	0.03	-0.13	0.000	18	0.02	-0.10	0.000	9	0.04	-0.08	0.000
Total Article Influence Score, previous 10 years	72	0.03	-0.18	0.000	18	0.01	-0.06	0.000	8	0.04	-0.08	0.000
A-journal articles, previous 10 years	6	0.05	-0.13	0.000	2	0.02	-0.11	0.000	1	0.02	-0.04	0.000
PhD students advised	-				5	0.03	-0.20	0.000	1	0.08	-0.15	0.000
PhD committees	-				25	0.05	-0.33	0.000	5	0.07	-0.13	0.000
Based in the South	0.28	0.01	-0.03	0.169	0.34	0.02	-0.12	0.000	0.36	0.05	-0.09	0.000
Observations	5,876				49,199				61,052			

Notes: The table provides descriptive information for the pool of eligible evaluators in qualification exams in Italy and in Spain. In Italy it includes only evaluators who are based in an Italian university. Columns 1, 5 and 9 report mean values for each corresponding variable and sample. In columns 2, 3, 6, 7, 10 and 11 variables have been normalized to have zero mean and unit variance for individuals within each field and rank. Columns 4, 8 and 12 report the p-value of a t-test of the difference in means between male and female eligible evaluators in the corresponding variable. *Article Influence Score* (AIS) is only available for candidates in science, technology, engineering, mathematics, medicine. Information on publications in *A-journal articles* is only provided for candidates in social sciences and humanities. In Italy, southern regions refer to Abruzzo, Molise, Campania, Apulia, Basilicata, Calabria and islands. In Spain, southern regions include Extremadura, Castille-La Mancha, Andalusia, Murcia, Valencia and islands.

TABLE C2—**Descriptive statistics – Applications**

	1	2	3	4	5	6	7	8	9	10
	Applications to full professorships					Applications to associate professorships				
	Mean	St.Dev.	Male	Female	p-value	Mean	St.Dev.	Male	Female	p-value
	Italy									
Female	0.31					0.41				
Age	49	8	-0.01	0.01	0.199	43	7	0.02	-0.03	0.000
Permanent position:	0.74	0.44	0.72	0.77	0.000	0.47	0.50	0.46	0.48	0.000
- same field	0.77	0.42	0.76	0.80	0.000	0.74	0.44	0.72	0.76	0.000
CV length (pages)	20	79	-0.01	0.03	0.006	14	60	-0.03	0.04	0.000
All Publications:	89	83	0.04	-0.09	0.000	53	54	0.04	-0.06	0.000
- Articles	53	65	0.06	-0.14	0.000	30	41	0.07	-0.10	0.000
- Books	3	6	0.04	-0.09	0.000	2	4	0.06	-0.08	0.000
- Book chapters	10	15	0.01	-0.03	0.004	6	10	0.01	-0.02	0.002
- Conference proceedings	14	26	-0.01	0.02	0.050	8	17	-0.01	0.01	0.028
- Patents	0.35	2.09	0.01	-0.03	0.000	0.19	1.39	0.03	-0.04	0.000
- Other	10	27	-0.00	0.00	0.938	8	20	-0.00	0.00	0.899
Number of coauthors per article	6	19	-0.01	0.02	0.024	6	18	-0.02	0.03	0.000
First-authored	0.22	0.19	-0.01	0.02	0.037	0.22	0.2	0.00	-0.00	0.365
Last-authored	0.15	0.17	0.02	-0.04	0.000	0.11	0.15	0.02	-0.03	0.000
Total AIS	90	127	0.05	-0.13	0.000	51	82	0.05	-0.08	0.000
A-journal articles	6	9	0.05	-0.08	0.000	3	6	0.04	-0.05	0.000
Application order	0.50	0.29	0.50	0.51	0.012	0.50	0.29	0.50	0.50	0.717
Above the median in 3 indicators	0.42	0.49	0.45	0.37	0.000	0.36	0.48	0.37	0.33	0.000
Withdrawal	0.16	0.37	0.15	0.20	0.000	0.13	0.34	0.12	0.16	0.000
Qualified	0.36	0.48	0.37	0.34	0.000	0.37	0.48	0.38	0.35	0.000
Failure	0.48	0.50	0.48	0.46	0.012	0.50	0.50	0.5	0.5	0.975
Proportion of positive votes	0.46	0.47	0.46	0.44	0.051	0.44	0.47	0.45	0.43	0.000
Number of applications	21,594					47,426				
	Spain									
Female	0.27					0.40				
Age	46	6	-0.01	0.03	0.015	37	6	0.03	-0.05	0.000
All Publications:	19	21	0.03	-0.09	0.000	8	14	0.07	-0.10	0.000
- Articles	17	21	0.04	-0.09	0.000	7	14	0.07	-0.11	0.000
- Books	0.64	1.47	0.01	-0.03	0.005	0.21	0.65	0.02	-0.02	0.000
- Book chapters	1.57	3.18	0.01	-0.02	0.086	0.54	1.41	0.01	-0.01	0.025
- Patents	0.04	0.33	-0.00	0.00	0.919	0.02	0.22	0.01	-0.01	0.012
Average number of coauthors	3	10	-0.00	0.01	0.691	5	23	-0.00	0.00	0.863
First-authored	0.25	0.31	0.00	-0.00	0.862	0.26	0.34	0.01	-0.01	0.200
Last-authored	0.24	0.30	0.01	-0.02	0.220	0.17	0.30	0.03	-0.05	0.000
Total AIS	18	27	0.01	-0.04	0.031	8	22	0.05	-0.11	0.000
A-journal articles	2	5	0.04	-0.09	0.000	1	2	0.06	-0.06	0.000
PhD students advised	2	3	0.03	-0.09	0.000	0.24	0.88	0.03	-0.05	0.000
PhD committees	7	9	0.03	-0.08	0.000	1	3	0.05	-0.08	0.000
Qualified	0.11	0.31	0.11	0.09	0.003	0.12	0.32	0.12	0.11	0.025
Number of applications	13,444					17,799				

Note: Columns 1 and 6 report mean values for each corresponding variable and sample. In columns 3, 4, 8, and 9 all productivity variables have been normalized to have zero mean and unit variance for applications within each exam. Columns 5 and 10 report the p-value of a t-test of the difference in means between male and female candidates in the corresponding variable. *Article Influence Score* (AIS) is only available for candidates in science, technology, engineering, mathematics and medicine. Information on publications in *A-journal articles* is only provided for candidates in social sciences and humanities.

TABLE C3—**Descriptive statistics – Links and Research Overlap**

	1	2	3	4	5
	All		Male	Female	
Italy	N	Mean	Mean	Mean	p-value
Colleagues	2,555,839	0.028	0.027	0.029	0.000
Coauthors	2,555,839	0.014	0.015	0.013	0.000
Same subfield	1,373,825	0.598	0.597	0.599	0.020
Spain					
Colleagues	5,445,067	0.045	0.047	0.043	0.000
Coauthors	5,445,067	0.004	0.0045	0.004	0.000
PhD advisor	5,445,067	0.002	0.002	0.002	0.322
PhD thesis committee	5,445,067	0.013	0.014	0.011	0.000
Overlap in research interests	4,711,621	0.196	0.183	0.218	0.000

Note: The table provides information on links between candidates and eligible evaluators within each discipline. Information about research interests is only available for candidates with a permanent contract in an Italian university and for candidates who have defended their thesis in Spain or who have participated in a thesis committee in Spain. The variable *Same subfield* takes value one if a candidate and an eligible evaluator belong to the same subfield (*settore scientifico-disciplinare*). The variable *Overlap in research interests* measures the degree of overlap between the research interests of eligible evaluators and candidates, as measured by their participation in PhD thesis committees. Column 5 reports the p-value of a t-test of the difference in means between male and female candidates in the corresponding variable.

Appendix D. First-stage estimates

Below we report the first-stage estimates from the IV estimations of the effect of committee gender composition on the relative success of female candidates reported in Table 1.

TABLE D1—**First-stage estimates**

Dependent variable:	$Female_e^{final}$	$Female_i * Female_e^{final}$	$Female_i * Female_e^{final}$
Second-stage estimates:	Column 4 of Table 1		Column 5 of Table 1
		Italy	
$Female_e^{initial}$	0.822	-0.006	
	(0.048)	(0.006)	
$Female_i * Female_e^{initial}$	-0.042	0.788	0.810
	(0.032)	(0.066)	(0.055)
Controls:			
$Female_e^{expected}$	Yes	Yes	Yes
$Female_i * Female_e^{expected}$	Yes	Yes	Yes
Candidate characteristics	Yes	Yes	Yes
Exam FE			Yes
F statistics:	188	74	218
Sanderson-Windmeijer F statistics:	323	380	218
		Spain	
$Female_e^{initial}$	0.954	0.000	
	(0.021)	(0.002)	
$Female_i * Female_e^{initial}$	0.014	0.970	0.966
	(0.012)	(0.017)	(0.020)
Controls:			
$Female_e^{expected}$	Yes	Yes	Yes
$Female_i * Female_e^{expected}$	Yes	Yes	Yes
Candidate characteristics	Yes	Yes	Yes
Exam FE			Yes
F statistics:	2396	3135	2310
Sanderson-Windmeijer F statistics:	2116	3503	2310

Note: Standard errors are clustered by committee.

Appendix E. Nonlinearities

The effect of the gender composition of committee on the relative success rate of females may be nonlinear for a number of reasons. First, the presence of a woman in the committee may affect the voting behavior of male evaluators (see section E of the paper). If this is the case, the transition from zero to one female evaluator in the committee may have a different effect than the transition from one to two female evaluators, or from two to three female evaluators. Second, decisions in the committee are taken on a (qualified) majority basis. Therefore, having a committee where the (qualifying) majority of members are female might have a particularly strong effect.

In order to correctly identify the potential existence of nonlinear effects, it is necessary to control for the probability that a given number of women is assigned to the committee. We consider the following model:

$$y_{ie} = \beta_0 + \beta_1 Female_i + \sum_k \gamma_k Female_i D_{ke} + \sum_k \delta_k Female_i D_{ke}^{expected} + \mathbf{X_i}\beta_2 + \mathbf{Z_i}\beta_3 + \mu_e + \epsilon_{ie} \quad (10)$$

where D_{ke} is a dummy variable that takes value one if the number of female evaluators in committee e is equal to k and $D_{ke}^{expected}$ is the probability that exactly k female evaluators are assigned to a given committee. For Spanish evaluations, we directly compute these probabilities using information on the gender mix of the pool of eligible evaluators. For the Italian case, the direct computation is more complicated, since the assignment procedure required no more than one committee member from each university. Instead, we compute these probabilities using the outcomes of 1,000,000 simulated random draws, which account for the restrictions on the randomization.

Committees rarely included more then three women. Therefore, we only analyze the effect of having one, two, and three or more female evaluators. In both countries, four positive votes are required for qualification. The estimation results are presented in Table E1. Overall, the linearity of the effect of committees' gender composition cannot be rejected by the data.

TABLE E1—**Nonlinearities**

	1	2
	Italy	**Spain**
Female	0.000	-0.012
	(0.007)	(0.007)
Female* 1 female evaluator	-0.017	-0.002
	(0.012)	(0.010)
Female* 2 female evaluators	-0.036	-0.005
	(0.012)	(0.013)
Female* 3 or more female evaluators	-0.079	-0.005
	(0.022)	(0.014)
Number of observations	69020	31243

Note: IV estimates. All regressions include as controls exams fixed-effects, the number of female evaluators in the committee, individual predetermined characteristics, and the expected probabilities to have 1, 2, and 3 or more female evaluators. Standard errors are clustered by committee.

Appendix F. Committee composition and evaluators' resignations

In section E of the paper we estimate the effect of committee gender composition on the voting behavior of male evaluators. The consistency of these estimates relies on the assumption that evaluators' resignation was not affected by the gender composition of the committee. To examine this possibility, we estimate the following equation on the sample of initially drawn evaluators:

$$\begin{aligned} Evaluator_{je}^{final} &= \beta_0 + \beta_1 Female_j + \beta_2 Female_e^{initial} + \beta_3 Female_j * Female_e^{initial} \\ &+ \beta_4 Female_{je}^{expected} + \beta_5 Female_j * Female_{je}^{expected} + \epsilon_{ij}, \end{aligned} \tag{11}$$

where $Evaluator_{je}^{final}$ is an indicator for those initially drawn evaluators who served in the final evaluation committee, $Female_e^{initial}$ is the share of women in the initially drawn committee and $Female_{je}^{expected}$ is the expected share of women in the committee conditional on the inclusion of evaluator j. Results from the estimation of equation (11) are reported in column 1 of Table F1. The presence of women in the committee does not affect the likelihood that a male or a female evaluator resigns. In column 2, we control for a number of evaluator characteristics including tenure, quality-adjusted productivity (total Article Influence Score in Sciences and the number of A-journal articles for Social Sciences and Humanities), and the location of their university. The estimates are unaffected by the inclusion of these controls.

TABLE F1—**The effect of committee composition on evaluators' resignations**

	Italy	
Female evaluator	0.117	0.106
	(0.093)	(0.096)
Share of women in the committee	0.118	0.115
	(0.071)	(0.071)
Female evaluator * Share of women in the committee	-0.085	-0.079
	(0.170)	(0.174)
Expected share of women	-0.187	-0.173
	(0.192)	(0.196)
Female evaluator * Expected share of women	-0.118	-0.127
	(0.289)	(0.293)
Controls:		
Evaluator characteristics		Yes
Mean dependent variable	0.922	0.922
Number of observations	920	920

Note: OLS estimates. The dependent variable is an indicator that takes value one if the initially drawn evaluator serves in the final evaluation committee. Standard errors are clustered by committee.

Appendix G. The effect of connections, by gender of evaluators and candidates

In Table G1 we explore whether the impact of connections depends on the gender of evaluators and candidates. We consider coauthors, colleagues and, in the case of Spain, also advisors. The gender of connections does not seem to play any role. Male and female candidates benefit equally from the presence of a female or a male connection in the committee.

TABLE G1—**The effect of strong connections, by candidate and evaluator gender**

	1	2
	Italy	**Spain**
Female candidate	0.006	-0.012
	(0.008)	(0.007)
Female candidate * Share of female evaluators	-0.131	-0.012
	(0.035)	(0.028)
Share of connections in committee	0.204	0.427
	(0.040)	(0.038)
Female candidate * Share of connections in committee	0.010	0.020
	(0.065)	(0.060)
Share of female connections in committee	-0.008	-0.036
	(0.085)	(0.101)
Female candidate * Share of female connections in committee	0.149	-0.084
	(0.125)	(0.145)
Number of observations	69020	31243

Note: IV estimates. All regressions include exam fixed-effects, individual predetermined characteristics, *Female candidate* Expected share of women in committee*, *Expected connections in committee*, *Female candidate* Expected connections in committee*, *Expected female connections in committee* and *Female candidate* Expected female connections in committee*. Standard errors are clustered by committee.

Appendix H. Heterogeneity analysis, alternative specifications.

In section F of the paper we explore whether the effect of the committees' gender composition varies depending on whether evaluators and candidates share similar research interests and depending on the degree of feminization of the field. In Table 7 we report results from an analysis where we split the sample of candidates in each country in two groups based on the median value of each variable.

In this section of the Appendix we present an alternative specification. We estimate a model with triple interactions exploiting the full range of possible values of the running variables. First, we analyze the impact of research similarity. We estimate the following model:

$$\begin{aligned} Y_{ie} = {} & \beta_0 + \beta_1 Female_i + \beta_2 Female_i * Female_e^{final} + \beta_3 S_{ie}^{final} + \beta_4 Female_i * S_{ie}^{final} + \\ & + \beta_5 Female_i * Female_e^{final} * S_{ie}^{final} + \\ & + \beta_6 Female_i * Female_e^{expected} + \beta_7 S_{ie}^{expected} + \beta_8 Female_i * S_{ie}^{expected} + \\ & + \beta_9 Female_i * Female_e^{final} * S_{ie}^{final} + \mathbf{X_i}\beta_{10} + \mu_e + \epsilon_{ie} \end{aligned} \tag{12}$$

where S_{ie}^{final} and $S_{ie}^{expected}$ stand for the actual and the expected research similarity between candidate i and committee e. We instrument the final composition of the committee ($Female_e^{final}$, S_{ie}^{final}) using the outcome of the initial lottery draw ($Female_e^{initial}$, $S_{ie}^{initial}$). In this model, coefficient β_2 shows the effect of committee gender composition when the committee members and candidates do not share research interests ($S_{ie}^{final} = 0$). Coefficient β_5 shows how this effect changes when candidates are evaluated by committees composed of evaluators who share research interests with the candidate.

As shown in Table H1, columns 1 and 2, in both countries the presence of women in the committee reduces the relative chances of success of female candidates when candidates and evaluators have different research interests. However, this effect disappears when candidates and evaluators share the same research interests.

Second, analyze the impact of the degree of feminization of the field. We estimate the following equation:

$$\begin{aligned} Y_{ie} = {} & \beta_0 + \beta_1 Female_i + \beta_2 Female_i * Female_e^{final} \\ & + \beta_3 Female_i * Female_e^{final} * Female_d + \beta_4 Female_i * Female_e^{expected} \\ & + \beta_5 Female_i * Female_e^{final} * Female_d + \mathbf{X_i}\beta_6 + \mu_e + \epsilon_{ie}, \end{aligned} \tag{13}$$

where $Female_d$ is the proportion of women among full professors in the corresponding discipline. Equation (13) allows us to explore whether the effect of committee gender composition varies depending on the feminization of the discipline. Results from this analysis are shown in columns 3 and 4 of Table H1. The effect of committees' gender composition on female candidates' success rate does not depend significantly on the degree of feminization of the field.

TABLE H1—**The effect of committee composition, by research interest overlap and the degree of feminization of the field**

	1	2	3	4
	Italy	**Spain**	**Italy**	**Spain**
Female	0.026	-0.004	0.010	-0.011
	(0.031)	(0.020)	(0.009)	(0.007)
Female * Share of women in the committee	-0.177	-0.103	-0.032	-0.008
	(0.077)	(0.047)	(0.109)	(0.048)
Research similarity	0.107	0.224		
	(0.049)	(0.047)		
Female * Research similarity	-0.072	-0.027		
	(0.063)	(0.072)		
Share of women in the committee * Research similarity	-0.216	-0.031		
	(0.138)	(0.093)		
Female * Share of women in the committee * Research similarity	0.242	0.308		
	(0.106)	(0.103)		
Female * Share of women in the committee * Feminization of the discipline			-0.350	-0.046
			(0.421)	(0.269)
Average research similarity	0.455	0.262		
Average feminization of the discipline			0.218	0.128
Candidate characteristics	Yes	Yes	Yes	Yes
Exam FE	Yes	Yes	Yes	Yes
Number of observations	35832	27998	69020	31243

Note: Characteristics of final committees are instrumented by the characteristics of initial committees selected by random draw. Research similarity is measured in Italy as the proportion of committee members in the same subfield as the candidate and in Spain as the average overlap in research interests (see more details in Data section). Feminization of the discipline is measured by the proportion of women among all full professors in the discipline in 2012 in Italy and in 2002 in Spain. Columns 1 and 2 include also *Female candidate* Expected share of women in committee*, *Expected research similarity*, *Female candidate* Expected research similarity*, *Expected share of women in committee * Expected research similarity* and *Female candidate* Expected share of women in committee * Expected research similarity*. Columns 3 and 4 include *Female candidate* Expected share of women in committee* and *Female candidate* Expected share of women in committee * Feminization of the discipline*. Standard errors are clustered by committee.

REFERENCES

Abrevaya, Jaso, and Daniel S. Hamermesh. 2012. "Charity and Favoritism in the Field: Are Female Economists Nicer (to Each Other)?" *Review of Economic and Statistics*, 94(1): 202–207.

Broder, Ivy E. 1993. "Review of NSF Economics Proposals: Gender and Institutional Patterns." *American Economic Review*, 83(4): 964–970.

Casadevall, Arturo, and Jo Handelsman. 2014. "The Presence of Female Conveners Correlates with a Higher Proportion of Female Speakers at Scientific Symposia." *mBio*, 5(1): e00846–13.

De Paola, Maria, and Vincenzo Scoppa. 2015. "Gender Discrimination and Evaluators' Gender: Evidence from the Italian Academy." *Economica*, 82(325): 162–188.

Ellemers, Naomi, Henriette Van den Heuvel, Dick de Gilder, Anne Maass, and Alessandra Bonvini. 2004. "The Underrepresentation of Women in Science: Differential Commitment or the Queen Bee Syndrome?" *British Journal of Social Psychology*, 43(3): 315–338.

Fuentes, Eulàlia, and Llorenç Arguimbau. 2010. "Las Tesis Doctorales en España (1997-2008): Análisis, Estadísticas y Repositorios Cooperativos." *Revista Española de Documentación Científica*, 33(1): 63–89.

Giménez-Toledo, Elea, Carlos Tejada-Artigas, and Jorge Mañana-Rodríguez. 2012. "Scholarly Publishers Indicators (SPI). 1st Edition."

Jayasinghe, Upali W., Herbert W. Marsh, and Nigel Bond. 2003. "A Multilevel Cross-Classified Modelling Approach To Peer Review Of Grant Proposals: The Effects Of Assessor And Researcher Attributes On Assessor Ratings." *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 166(3): 279–300.

Maraut, Stéphane, and Catalina Martínez. 2014. "Identifying Author-Inventors from Spain: Methods and a First Insight into Results." *Scientometrics*, 101(1): 445–476.

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*, 100(6): 1678–1712.

Moss-Racusin, Corinne A., John F. Dovidio, Victoria L. Brescoll, Graham Mark J., and Jo Handelsman. 2012. "Science Faculty's Subtle Gender Biases Favor Male Students." *Proceedings of the National Academy of Sciences*, 109(41): 6474–16479.

Steinpreis, Rhea E., Katie A. Anders, and Dawn Ritzke. 1999. "The Impact of Gender on the Review of the Curricula Vitae of Job Applicants and Tenure Candidates: A National Empirical Study." *Sex Roles*, 41(7/8): 509–528.

Torres-Salinas, Daniel, Maria Bordons, Elea Giménez-Toledo, Emilio Delgado-López-Cózar, Evaristo Jiménez-Contreras, and Elías Sanz-Casado. 2010. "Clasificación integrada de revistas científicas (CIRC): propuesta de categorización de las revistas en ciencias sociales y humanas." *El profesional de la información*, 19(6): 675–683.

Williams, Wendy M., and Stephen J. Ceci. 2015. "National Hiring Experiments Reveal 2:1 Faculty Preference for Women on Stem Tenure Track." *Proceedings of the National Academy of Sciences*, 112(17): 5360–5365.