# Gender Quotas in Hiring Committees: a Boon or a Bane for Women?

Pierre Deschamps

SciencesPo
LABORATOIRE INTERDISCIPLINAIRE
D'ÉVALUATION DES POLITIQUES PUBLIQUES

LIEPP Working Paper

**Novembre 2018, nº82**

# Gender Quotas in Hiring Committees: a Boon or a Bane for Women?

**Pierre DESCHAMPS**
SOFI, University of Stockholm and LIEPP
pierre.deschamps@sofi.su.se

www.sciencespo.fr/liepp


Comment citer cette publication:
Pierre Deschamps, Gender Quotas in Hiring Committees: a Boon or a Bane for Women?, *Sciences Po LIEPP Working Paper* n°82, 2018-11-29 (last update 2020-07-06).

# Gender Quotas in Hiring Committees: a Boon or a Bane for Women?*

Pierre Deschamps[†]

June 19, 2020

**Abstract**

Women are underrepresented in academia. In order to reduce hiring discrimination, the French government imposed a gender quota in academic hiring committees in 2015. Drawing on a unique dataset provided by French universities together with a difference-in-difference design, I estimate the causal effect of the reform on the probability of women being hired. I exploit the reform's 40% threshold by assigning fields in universities whose committees were on average below the threshold to the treatment group, and those that were already respecting the quota to the control group. I show the reform backfired and significantly worsened both the ranking of women and their probability of being hired, with a treatment effect comparable to a 8% decrease in the share of women recruited. Since the negative effect of the reform is concentrated in committees that are helmed by men, this result seems to be driven by the reaction of men to the reform.

(JEL J16, J71)

## 1 Introduction

Even though women make up the majority of PhD candidates in many academic disciplines, they are under-represented in faculty positions, and the gender gap increases with seniority.[1] Beyond this vertical segregation, academia also suffers from horizontal segregation, with women much less likely to be studying and working in STEM. Resolving this

*I am very grateful to Etienne Wasmer for his advice and help in starting this project. I also thank Mathieu Arbogast, Ghazala Azmat, Manuel Bagues, Thomas Chaney, Alain Chenu, Etienne Fize, Daniel Hamermesh, Emeric Henry, Nicolas Jacquemet, Morgane Laouenan, Erik Lindqvist, Thierry Mayer, José Montalban, Roland Rathelot, Anne Revillard, Jean-Marc Robin, Dan-Olof Rooth, Marie Sautier and Jose de Sousa for their constructive comments, as well as seminar participants at Sciences Po, SOFI, the Banque de France EPP conference, and the SOLE, AFSE, EEA, RES PhD and EALE meetings. I thank Michel Gardette, Isabelle Delacroix, Magali Steinmann, Marion Paoletti, Arnaud Dudek, Bruno Oliveira and Anne Kupiec for their help in accessing the data. This work is supported by a public grant overseen by the French National Research Agency (ANR) as part of the "Investissements d'Avenir" program within the framework of the LIEPP center of excellence (ANR11LABX0091, ANR 11 IDEX000502).

†Pierre Deschamps: SOFI, University of Stockholm, pierre.deschamps@sofi.su.se

[1] In the European Union for instance, women made up 47% of PhD graduates, 40% of assistant professors, but only 23% of full professors, according to the European Commission's 2018 SHE figures.

disparity could lead to important aggregate productivity gains in research.[2]

There are many potential causes for this gender gap, such as discrimination against women, differences in productivity, which could be caused by the unequal distribution of household work and childcare (Goldin, 2014; Blau and Kahn, 2017), or differences in attitudes to risk and competition (Niederle and Vesterlund, 2007). It is unclear from the literature which cause is the most salient. If hiring discrimination from male professors against their female colleagues is one of the main causes, then gender quotas on the members of hiring committees could be a solution to the under-representation of women. Quotas can also come with significant costs however, such as an increase in administrative workload for female professors or a potential decrease in the "quality" of hiring committees and thus recruitment. The success of gender quotas in hiring committees therefore depends on whether they improve the probability of women being hired. Though previous studies have leveraged random variation in the gender composition of hiring committees, none have directly estimated the effect of these gender quotas on the hiring of female professors.

To estimate the causal effect of gender quotas in hiring committees, I take advantage of a reform by the French government. In order to reduce gender bias in recruitment from male professors,[3] the reform stated that from 2015 onwards, academic hiring committees in the public sector would have to be made up of at least 40% of members of each gender.[4] This feature makes it possible to estimate the causal effect of the reform by employing a difference-in-difference design using pre-reform distance to the 40% threshold. Fields in universities where hiring committees were below the 40% threshold on average are assigned to the treatment group, and the remaining observations to the control group.

Data on the gender composition of committees and the composition of the applicant pool are not publicly available. I rely instead on data collected by hand from a set of university archives, from which I received access to the names of committee members, the names of ranked candidates, the composition of the applicant pool, the gender of the jury president, the most powerful position in the committee, and the final ranking given by

[2]For instance, Hsieh et al. (2019) argue that a better allocation of talent due to decreasing barriers to entry accounts for 40 percent of US GDP growth between 1960 and 2000.

[3]This reform was proposed in the bill under the heading "Measures on equal opportunities between men and women, and reducing discrimination". Law N°2012-347, Article 55.

[4]Decree n°84-431

the committee, which determines which candidates are eventually hired.

I find that the quota has a large effect on the treatment group, where the average share of women in committees moves from 30% to 45%. As expected, the reform has no effect in the control group. This variation is then used to study the effect of the reform on the ranks women receive, and on the probability women are hired by the university, as well as the effect of the quota on the gender of jury presidents and average qualifications of the committee.

First, I find that the quota has a strong negative effect on the hiring and ranking of women by these committees. The treatment effect is equivalent to an 8% (4 percentage points) decrease in the share of women recruited. This result holds across a plethora of different specifications (continuous versus binary treatment variable, IV) and outcomes (hiring, ranks, dyads). The effect is not driven by inattention of jury members to lower ranked candidates, or longer-run time-trends.

Second, I use the same assignment to treatment and control groups to evaluate the effect of the quota on the share of female jury presidents, the average h-index of the jury and the gender composition of the application pool.[5] This allows me to evaluate the effect of the quota on the representation of women in academic decision-making,[6] the average "quality" of committees, and the effect of role-models on the probability of women staying in academia, respectively. Though the share of women in recruitment committees moves from 40 to 47% in my sample overall, there is no increase in the share of women elected jury presidents, the key position in these committees. There doesn't seem to be any effect on the gender composition of the applicant pool either, at least in the short-term context of this paper. I find no change in the average quality of the jury following the reform, as proxied by the average h-index.[7]

I control for publications and connections in my analysis, using data scraped from online sources to investigate concerns that women are positively or negatively selected.

[5] The h-index of jury members is the $h$ number of publications with more than $h$ citations.

[6] See Azmat (2014) for positive and negative aspects of increasing gender representation in decision-making teams.

[7] Gender quotas also come with significant costs for women. The relatively small number of women eligible to participate in committees imply that gender quotas in committees will substantially increase their administrative workload, potentially harming their ability to publish. One result from this article is that most of the new female members on the committees are internal members, indicating that women have to do more administrative work for their department following the reform.

The main results vary little when these control variables are added. I find that having more citations has a positive effect on your ranking, but so does having an advisor in the committee, or having a PhD from the university to which you are applying. Men seem to benefit more from having a supervisor in the committee than women do, confirming results from qualitative research.

What is the cause of this surprising negative effect of quotas on the hiring of women? Either women have opposite-gender biases, or men change their behaviour as a result of the reform. I find that the negative effect of the reform is entirely driven by committees that are helmed by men. Jury member fixed effects also show the same pattern. This suggests that changes in the gender-bias of men drive the negative effect of the reform.

This paper is the first to directly estimate the effect of gender quotas in hiring committees on the recruitment of women. Previous articles have indirectly inferred the effects of gender quotas from random assignment of women to committees. Bagues and Esteve-Volart (2010) and Bagues et al. (2017) find that women are less likely to be hired when more women are randomly assigned to a candidate's promotion committee, results which are similar to the ones in this paper.[8] If random assignment increases or decreases biases,[9] directly estimating the impact of the quota will provide more precise evidence on the effects of gender quotas on the hiring of women. This distinction is relevant from a policy standpoint, given that there is no random assignment of recruiters to candidates in most recruiting contexts. A second contribution relative to these papers is the estimation of the effect of quotas on several other dimensions such as the reaction of the gender composition of applicant pools to the quota and the lack of change in the quality of jury members. These dimensions are necessary to understand how gender quotas change the hiring process.[10] Though the methodologies and objects studied are different, my paper and these articles imply that increasing the share of women in the recruitment process is unlikely to solve the under-representation of women in academia.

The results cast doubts on the hypothesised "trickle-down" benefits of gender quotas

[8]Paola and Scoppa (2015) find positive effects on the hiring of women, but only when the number of women in the committee increases from 0 to 1.

[9]See Li (2017), for evidence of a bias/efficiency trade-off when considering the assignment of candidates to jurors in scientific funding.

[10]Another important difference is the nature of the data. Both these papers analyze promotion decisions, where only a few marginal candidates might be affected by discrimination. Studying ranks allows for estimation of discrimination parameters over the whole distribution of candidate quality.

on women that do not directly benefit from them. The literature finds that boardroom quotas have little benefits for women lower down in the wage distribution (Bertrand et al., 2018).[11] The results of this article show that increasing female representation at the top can deteriorate outcomes for women on the labour market.

The proposed mechanism also sheds light on the literature that estimates own or opposite gender effects. Previous articles have found conflicting results. Some find that female evaluators have a positive effect on the outcomes for women (Lincoln et al., 2012; Boring, 2017; Zeltzer, 2020; Edo et al., 2015).[12] A few find no effect (Abrevaya and Hamermesh, 2012; Card et al., 2019; Feld et al., 2016; Williams and Ceci, 2015).[13] Ellemers *et al.* (2004), and Broder (1993), find opposite-gender preferences, in the evaluation of doctoral students' work ethics and grant applications (using evaluations from men and women on the same article) respectively. The wide disparity of estimated effects can be rationalised if gender preferences are context-dependent as suggested by Akerlof and Kranton (2000) and this paper.

In section 2, I explain how I construct the data set, and present some descriptive statistics. In section 3, I discuss how own- or opposite-gender preferences could intervene in the ranking of committees, and how we can estimate them. In section 4, I look at how the quota affected hiring, and discuss the interpretation of these results and potential channels in section 5. I conclude in section 6.

[11] Bertrand et al. (2018) find that boardroom quotas in Norway have a positive effect on the wages of women in the boardroom, and that these women are more qualified than before, but that these gains do not trickle-down the wage ladder. Maida and Weber (2019) and Dalvit et al. (2018) find similar effects of boardroom quotas for Italy and France.

[12] Lincoln et al. (2012) finds that having a male committee chair for research prizes raises the probability of men receiving a prize. Boring (2017) shows that male students give higher evaluation grades to their male professors. Zeltzer (2020) finds that male physicians refer patients to male specialists relatively more often than do female doctors. Edo et al. (2015) finds that female recruiters are more likely to call back women.

[13] Abrevaya and Hamermesh (2012) and Card et al. (2019) look at whether female referees accept articles written by women more often and find no effect. Feld et al. (2016) find no own-gender preferences in graders' evaluation of students. Williams and Ceci (2015) run lab experiments, and find that women and men prefer to hire women over men with identical profiles, but that these preferences do not change depending on the evaluator's gender.

# 2 Data

In this section, I present the dataset compiled from administrative data on 455 hiring committees from 3 different French universities. The data spans from 2009 to 2018, but the bulk of the data is from committees post-2013, as can be seen in Table A.4.

## 2.1 French hiring committees

French academic hiring committees are created *ad hoc* for each position that has to be filled. These committees have a jury president, who has broad powers over the nomination of committee members. The president also has a deciding vote in case of a tie between candidates. There is some variation in the number of members in these committees; by law there must be between 8 and 20 jurors. At least half of the members must be from outside the hiring university. Once a committee has been created, candidates can apply via a web platform called GALAXIE,[14] and post their CVs. In some cases (but not all), they are aware of the jury composition at the time of their application.[15]

As part of the French Government's push for greater gender equality, since the 1st January 2015, each committee must also be made up of at least 40%[16] of each gender. However, there are neither constraints on the gender of the jury president, nor quotas on the gender of the candidates hired.

The committee then decides which candidates can receive an on-campus job interview, and after these auditions, ranks the candidates it deems worthy (if any) of being hired by the committee. The committee votes on the final ranking, and not on the candidates individually. Though members can dissent during the final vote, the final ranking is almost always accepted unanimously. Because the individual behaviour of committee members is not observed, I can only infer the effect of increasing the number of women in committees from two variables: the gender of the jury president, and the percentage of women in the

[14] Link to GALAXIE and available positions.

[15] We could be worried by auto-censure mechanisms, although if these mechanisms exist, they would reinforce our results. Paola and Scoppa (2015) find that women tend to drop out when applying for positions more often than men, but this effect is independent of the composition of the committee.

[16] For maths, this ratio was dropped exceptionally to 14% and to 33% for Political Science until 2016 (Link to the decree). In other fields, this ratio is binding but once the committee has been approved, if a committee member drops out, the replacements do not need to respect the gender quota. The data includes the final composition of the committee.

hiring committee.

Offers are given sequentially in the order that candidates are ranked. All ranked candidates can potentially receive an offer from the university, if the candidates ranked above them refuse the offer.[17] The candidates are aware of the ranks that they receive, which helps them to inform their decision on which offers they should accept or reject. Ranks may also give a signal to candidates on whether it is relevant to pursue an academic career. This is important in our context since Geuna and Shibayama (2015) show that women are more likely to drop out of academia than men.

For each committee, I have access to the names and ages of the jurors and the names and ranks of the candidates, as well as the gender composition of the candidate pool in 85% of the cases. During the analysis, the ranking variable includes the K-ranked candidates, and the candidates that audition but are not ranked, who are all given the rank K+1. Since I only use order and not distance in my regressions, this normalisation is harmless. Candidates that are not auditioned are discarded for the regression analysis, but are included when computing the gender ratio of applicants.

## 2.2 Descriptive Statistics

I present some statistics on the committees in Table 1. As we can see, the number of female presidents and the share of women in committees varies strongly between fields. In some fields, female presidents are few and far between: Only 4 out of 28 recruitments in Physics for instance, or 2 in economics. In general, the percentage of women hired is correlated with the percentage of female presidents and the share of women in committees at the field level, but also with the percentage of female applicants.

For a small (72) number of contests, I am missing data on applicants, and use instead the percentage of women among the ranked candidates. The results from the main specifications are quantitatively similar when these observations are excluded. As shown in the columns of Table A.1 in the Appendix, which excludes these committees, the average share of female applicants is also a good predictor of average percentage of women ranked, especially when we consider only the fields for which we can observe more than

[17] This is relative infrequent in the data set. Only 20 first-ranked candidates refuse the position.

Table 1: Descriptive statistics on recruitment

| | # recruitments | # recruitments pre-reform | # recruitments post-reform | Share of women hired | Share of female presidents | Share of women jurors | Share of female applicants |
|---|---|---|---|---|---|---|---|
| Biology | 65 | 28 | 37 | 0.51 | 0.38 | 0.48 | 0.54 |
| Business | 18 | 4 | 14 | 0.33 | 0.22 | 0.49 | 0.42 |
| Chemistry | 18 | 5 | 13 | 0.39 | 0.44 | 0.44 | 0.39 |
| Economics | 27 | 11 | 16 | 0.19 | 0.07 | 0.41 | 0.26 |
| Education | 18 | 4 | 14 | 0.50 | 0.17 | 0.46 | 0.63 |
| Engineering | 18 | 5 | 13 | 0.06 | 0.11 | 0.40 | 0.23 |
| History | 12 | 10 | 2 | 0.33 | 0.42 | 0.37 | 0.28 |
| Languages | 40 | 16 | 24 | 0.47 | 0.47 | 0.52 | 0.56 |
| Law | 33 | 14 | 19 | 0.42 | 0.09 | 0.43 | 0.46 |
| Maths | 57 | 16 | 41 | 0.14 | 0.30 | 0.38 | 0.19 |
| Pharmacology | 30 | 9 | 21 | 0.57 | 0.40 | 0.48 | 0.53 |
| Physical Education | 13 | 7 | 6 | 0.15 | 0.08 | 0.46 | 0.15 |
| Physics | 28 | 12 | 16 | 0.21 | 0.14 | 0.40 | 0.20 |
| Political Science | 19 | 12 | 7 | 0.26 | 0.32 | 0.38 | 0.35 |
| Psychology | 30 | 14 | 16 | 0.57 | 0.43 | 0.50 | 0.62 |
| Sociology | 29 | 21 | 8 | 0.45 | 0.40 | 0.37 | 0.43 |
| Total | 455 | 188 | 267 | 0.36 | 0.30 | 0.44 | 0.40 |

The statistics above are compiled on a dataset that includes assistant and full professors, over the years 2009-2018.

20 committees. An important takeaway from Table 1 is that most of the between-field difference in the hiring of women is driven by differences in the composition of the applicant pool.[18] This is consistent with another article by Bosquet et al. (2018) that also uses French data, and concludes that gender differences in promotion rates are mostly driven by gender differences in applications. This can be more readily seen from two graphs in Figure 1 that show the mean rates of hiring by field against the pre-reform share of women in committees and the percentage of female applicants per field. There is a strong positive correlation between both these variables and the share of women hired, but the relationship between female applicants and the gender of winners is much stronger.[19]

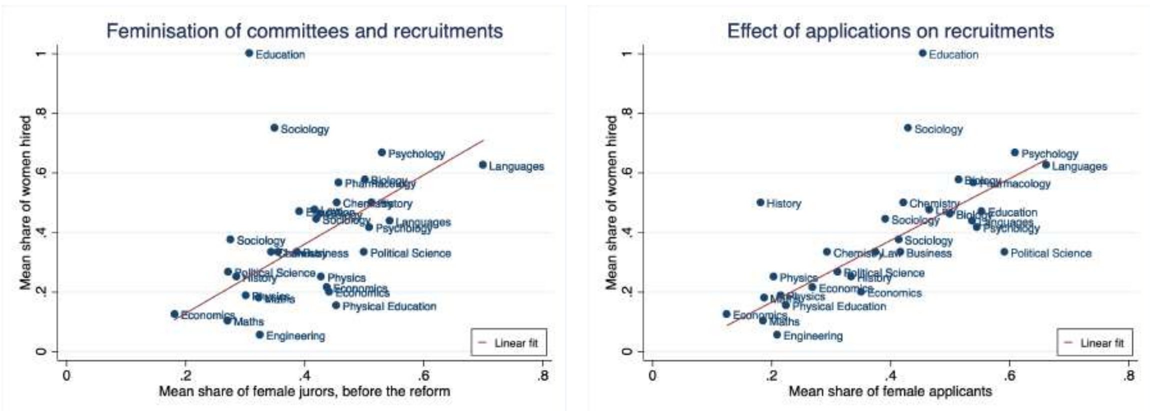



Figure 1: Determinants of hiring

[18]I cannot rule out that the coefficient from regressing whether a woman was ranked first on the share of women applicants or ranked is 1. This implies that the lack of women in STEM for instance is not driven by differential discrimination at the hiring stage relative to other fields, but rather by differences in the applicant pool.

[19]In Table B.5 which directly tests for the effect of these two variables, the positive effect of the feminisation of committees becomes insignificant once we control for percentage of female applicants.

Next, we can look at the effect of the reform on the share of women in committees. This reform affected fields heterogeneously. I present the average female/male juror ratio in Table A.3 for each discipline before and after the reform. We can see a clear increase in the proportion of women sitting on committees due to the reform, except for the disciplines where the quotas were already respected, such as Biology or Languages. There is also some variation across universities, since labs differ by feminisation.

When constructing the treatment and control groups, I use variation at the field-university level, assigning to the treatment group the disciplines in universities where the average proportion of women in committees pre-reform was less than 0.4. The variation at the field university level is shown graphically in Figure 2. All observations where the circle is to the left of the 0.4 threshold, represented by the vertical line, are assigned to the treatment group, while the others are assigned to the control group. As is clear from the Figure, the reform has a significant effect on the treatment group; the mean share of women in committees moves from 30% to 45%, with a t-stat of 9.89. There is no significant effect on the control group, where the mean share moves from 0.480 to 0.482 (t-stat 0.2).

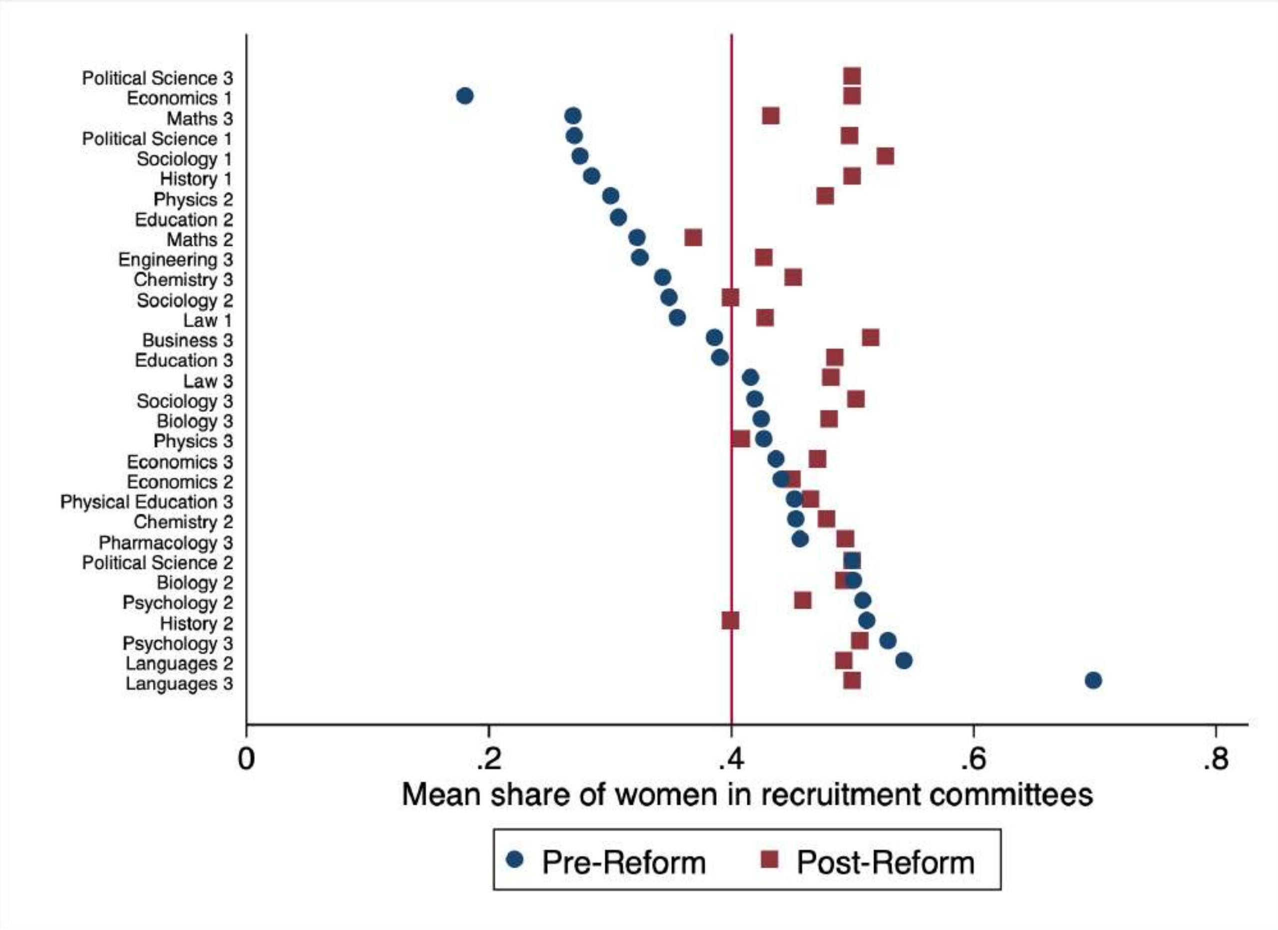
Though the reform has a clear effect on the share of women in committees, there is no systematic effect on the gender of the jury president however. This is similar to results in other contexts, which find that gender quotas do not necessarily help women gain access to the most prestigious positions. For instance, following the implementation of boardroom quotas in France, Rebérioux and Roudaut (2017) find that the newly promoted women are less likely to hold key positions in these boardrooms than newly promoted men.

### 2.2.1 Additional dimensions

Although I do not observe the CVs of candidates, I can control for two of the most important components of candidate quality, publications and connections. To control for publications, I collect the h-index and citations of candidates using the Scholar H-index calculator (2010), which creates h-indices and citation counts from Google Scholar.[20]

[20] A scientist's h-index is the $h$ number of publications he has over $h$ citations. I also collect citation and h-indices which are discounted by the number of co-authors and how long ago the article was published. Citations were collected in 2018 for articles published the year the committee convened.

Figure 2: Effect of the reform on the share of women in recruitment committees




Each entry represents a discipline (e.g. Maths, Law, Biology) in a university (1,2,3).

These variables will be used as measures, albeit imperfect,[21] of candidate quality. This matters if, on average, ranked female candidates have smaller or larger h-indices than male candidates. Table H.17 in the Appendix explores this possibility.

Academic connections are another potential confounder (see Combes et al. (2008)). In order to get more information on connections between jurors and candidates, I scraped the French website, theses.fr, which gathers information on which institution French doctors received their PhD from, as well as the name of their PhD supervisor. I collected similar data for foreign PhD students from other online sources.[22]

As we can see in Table A.2, the probability of having one's PhD supervisor in the committee is small (6%) but not negligible, rising to 12% in Psychology. We can also

[21] As pointed out by Bagues and Zinovyeva (2015), uninformed jurors are likely to only consider citation or h-index measures of candidates to determine whether the candidates are qualified or not, whereas more informed jurors will rely on information that can not be observed.

[22] Most juries are made up almost exclusively of French professors, so the probability of students who have received a PhD abroad of being in a committee with their PhD supervisor is very small.

look at endogamy, how often candidates with a PhD from the same institute apply for a position there. Once again, this is highly dependent on fields, with candidates in Biology (around 39%) very likely to have received a PhD from the same institution where they are applying for a position.

# 3 Model

In the administrative data used in the article, I observe a single ranking from each committee over all candidates. I consider that jury $j$ has a latent utility function $U_{ij}$ when hiring candidate $i$ which is defined as follows:

$$U_{ij} = q_{ij} + \mu g_i g_j$$

where $q_{ij}$ is how qualified the candidate is for a particular post, as evaluated by the committee j, and $\mu$ is a parameter that evaluates how committees rank different genders, when $g_i$ is the gender of the candidate, and $g_j$ the gender composition of the committee.[23] In this article I consider two different methods to recover the parameter of interest, $\mu$, from ranked data: a rank-ordered logit and a conditional logit. Although the rank-ordered logit directly takes ranks into account, there is no clear method for performing IV, which is why I consider a random utility model using dyads in Appendix D.

## 3.1 Rank-ordered Logit

One way to estimate this model could be to regress the rank of the candidate using OLS on the independent variables. However, ranking is simply a preference ordering, and the distance between the $1^{st}$ and $2^{nd}$ rank might not be the same as between $2^{nd}$ and $3^{rd}$. The standard method to recover parameters from ranking data in economics is the rank-ordered logit, which doesn't rely on the distance between ranks.

This method can be viewed as a multiplicative form of the traditional multinomial logit, and was proposed by Beggs et al. (1981), later refined in Hausman and Ruud (1987). The intuition is that respondents rank items based on repeated multinomial logit

[23] Since we have ranking data, I do not consider intransitive preferences, and do not study heterogeneity of utilities among committee members.

decisions, first choosing item $i$ from $M$ choices, then item $i'$ from $M-1$ choices and so forth, with the choice set decreasing in size until only one alternative is left. Combes et al. (2008), use this method to analyse whether connections matter for academic promotions, using ranking data from the French *agrégation* in economics, a context very similar to ours.

Putting more structure on $U_{ij}$, consider the case where $q_{ij}$ has the following form:

$$q_{ij} = x_{ij}\beta + \eta_j + \epsilon_{ij}$$

where $x_{ij}$ is a vector with the characteristics of the candidates, $\eta_j$ is a jury fixed effect and $\epsilon_{ij}$ is an error term. Consider two variables, $q^*_{ij}$, which is equal to the rank $k$ given by jury j to alternative i, and $r_{ij}$, which is equal to the alternative $i$ given the rank $k$ by jury j. If $\epsilon_{ij}$ is distributed Type 1 extreme value, so that $Pr(\epsilon_{ij} < u) = e^{-e^{-u}}$, then the likelihood of observing a particular ranking in the $j^{th}$ concours is:

$$Pr(U_{r_{1j}j} > U_{r_{2j}j} > \ldots > U_{r_{Mj}j})$$

$$l_j(\beta) = \prod_{i:q^*_{ij}=1}^{M-1} \frac{\exp(x_{ij}\beta + \mu g_i g_j)}{\sum_{i':ri'j \geq rij}^{M} \exp(x_{i'j}\beta + \mu g_{i'} g_j)}$$

Notice here that the jury fixed effects, $\eta_j$ cannot be recovered, since they do not enter the likelihood function. We can only recover parameters that vary for candidates within jury. The log-likelihood function of J independent committees is then:

$$\sum_{j=1}^{J} \sum_{i:q^*_{ij}=1}^{M-1} x_{ij}\beta + \mu g_i g_j - \sum_{j=1}^{J} \sum_{i:q^*_{ij}=1}^{M-1} (\log \sum_{i':ri'j \geq rij}^{M} \exp(x_{i'j}\beta + \mu g_{i'} g_j))$$

The model can be estimated through maximum likelihood. Since the log likelihood is globally concave, $\beta$ and $\mu$ will be the unique maximisers. We can recover the parameters $\mu$ by considering (for instance) the effect on the rank of women (relative to men) of having a female president (or a committee made up of many female members) or the effect for both genders of having a president of the same gender, as estimated directly in Section B of the Appendix. The main estimation relies on variation from the quota reform. $\mu$ can be inferred from the differential effect on the quota on the treatment and control groups.

Instead of considering the direct, endogenous effect of $g_{i'}g_j$, I infer the size of $\mu$ using variation from the policy reform, replacing $g_j$ by a treatment group dummy variable $T_j$, a post reform dummy variable $R_j$, and the interaction between these two terms.

In the estimation, I also use a conditional logit to estimate the probability of being first ranked. The reasoning is very much the same as the one above, except that the conditional logit can answer the following question: What is the probability of ranking candidate i first from a pool of M candidates?

# 4 Effect of the quota

We can directly estimate the link between the share of women in committees and the rank of women. Results of these regressions can be seen in Section B of the Appendix. However, these are only correlational results. I test for endogeneity between the share of women in committees and the gender composition of applicants in Section B.1, and find evidence that this is a concern in my dataset. This suggests that there could also be a correlation between the unobserved characteristics of candidates and the share of women in committees. I rely instead on the quota to generate variation in the gender composition of these committees.

I use the reform to construct a control and a treatment group: fields in universities where the gender composition of the jury is below 40% before the reform in 2015 are assigned to the treatment group. For instance, if on average, committees in Law in University A are made up of 35% of women on average before 2015, and of 45% on average in University B, then committees in Law in University A will be assigned to the treatment group, and Law committees in University B will be assigned to the control group. In practice, most fields are consistently assigned to treatment or control groups across universities.

I first analyse the effect of the reform on ranks, since this is the specification that uses all the variation in the data. To test for the effect of the reform, I estimate the equation below:

$$\text{Rank*}_{ij} = \alpha R_j T_j g_i + \gamma T_j g_i + \delta R_j g_i + \beta_1 X_i + \beta_2 g_i Y_j + \epsilon_{ij}$$

where Rank* is the latent variable that the members of the jury use to rank candidates, $\gamma$ represents the effect of being in the treatment group before the reform, $\delta$ the effect of the reform on the control group, and $\alpha$ is the parameter of interest, the effect of being in the treatment group after the reform relative to the control group. $g_i$ is a binary variable equal to 1 if the candidate is a woman, $X_i$ is a vector of individual characteristics of the candidate (such as the candidate's h-index[24] or whether the candidate received his PhD from the university he is applying to), and $Y_j$ a vector of committee characteristics (such as the share of female applicants or a dummy for whether the post is for an assistant or associate professorship)[25] that are interacted with $g_i$. The model is estimated through rank-ordered logit.

The results using the rank-ordered logit are presented in Table 2. For all variables estimated at the committee level, the results presented are the differential effects for women with respect to men. As we can see, there is a strong negative effect on the rank of women of being in a field that had a large increase in the share of female jurors due to the reform.[26] There is no significant effect on the control group in this specification. This effect is robust to the inclusion of other covariates, such as the standardised h-index of candidates, having a supervisor in the committee or the effect of having a PhD from the same institute. To interpret the size of the effect, one can compare the coefficient in the first row with the one for the standardised h-index. In column 2 for instance, the size of the effect is equivalent to a 3 standard deviation drop in the h-index of candidates.[27]

The control variables here are significant and have the expected sign: having a supervisor in the committee, or having a PhD from the institute where you are applying increases your ranking by the committee, and the effect size here is larger than a 1 standard deviation increase in the h-index. This result confirms previous research on the importance of networks and connections in academia.[28]

[24]I use the age-adjusted h-index of candidates since this is the publication measure that is the most significant when including all other measures, as seen in Table H.20 in the Appendix. This measure is also standardised by field and seniority.

[25]Maitre de Conférences or Professeur des Universités in French.

[26]The standard errors presented are not clustered at field level. In fact clustering at discipline level or the discipline-university level actually lowers the standard errors for almost all the estimations, and the results in column 6 of Table 6 for instance become significant at the 5% level when clustering the standard errors.

[27]I find no evidence of different pre-trends, as can be seen in Table H.24 in the Appendix, though the limited time span of my data means this test has very little power.

[28]See Combes et al. (2008) or Colussi (2018) for examples of connections influencing promotions and

Another way to interpret this result is to count the women that weren't hired as a result of the reform, which we can recover using some back-of-the-envelope computations. I can simulate the predicted probabilities from the rank-ordered logit to estimate a predicted rank, and in particular a predicted first-ranked candidate. We can then analyse the effect of the quota under two counterfactual scenarios: One in which there is no effect of the reform in the treatment group (i.e. $\alpha = 0$ and $\delta$=0 if the contest is in the treatment group), and one in which there is no reform at all (i.e. $\alpha$ and $\delta$=0). These scenarios are presented in Table 3. When compared to the simulated number of women recruited, if there had been no effect of the reform on the treated, the number of women recruited would have increased by 24% (11 p.p). If there had been no effect of the reform on the treated and control groups, we would still have expected an increase of 8% (4 p.p) in the number of women recruited.[29]

The results can be considered as a Local Average Treatment Effect (LATE) i.e. the effect of increasing the share of female jurors in disciplines that have few women. These effects may be stronger or weaker than the ones we would find when analysing all fields. However, the measure that we are considering is the relevant one, since the policy proposals that are usually debated aim at increasing the proportion of women in fields where they are not represented, not in fields where parity is already achieved. Whether increasing the proportion of female evaluators also improves the rank of women in fields where parity is already respected is a moot point. Though the counterfactual estimation suggests that the net effect is negative, the estimated global effect of the reform, presented in Appendix E, is not significantly different from 0.

## 4.1 Results on first-ranked candidates

We can next consider whether the reform had any effect on first-ranked candidates. Since in our dataset, most candidates who are ranked first accept the offer, this tests directly for

publications.

[29]We could also be interested in interpreting the effect of the reform in terms of the increase in the share of women in recruitment committees. This would require instrumenting the share of women in recruitment committees with the effect of the quota. I propose two methods in the Appendix, one using dyads in Section D, and using a control function in the rank-ordered logit in Table F.13. These methods give the same conclusion: Increasing the share of women in committees by 10% has the same effect on women as a 2 standard deviation drop in the h-index.

Table 2: Effect of the quota on the rank of female candidates

| Dependent variable: | Rank | | |
|---|---|---|---|
| Estimator: | Rank-ordered logit | | |
| | (1) | (2) | (3) |
| *Interacted with gender of the candidate* | | | |
| **Post Reform × Treatment Group** | -0.639** | -0.962*** | -0.818** |
| | (0.31) | (0.37) | (0.40) |
| Post-Reform | 0.213 | 0.226 | 0.167 |
| | (0.20) | (0.21) | (0.23) |
| Treatment Group | 0.331 | 0.674** | 1.503** |
| | (0.22) | (0.32) | (0.59) |
| Female president | -0.069 | -0.045 | -0.114 |
| | (0.17) | (0.16) | (0.18) |
| Share of female candidates | | -0.186 | -0.245 |
| | | (0.40) | (0.43) |
| Professorship | | | 0.342* |
| | | | (0.19) |
| *Variables at the candidate level* | | | |
| Candidate is a woman | -0.323** | | |
| | (0.14) | | |
| Standardised age-adjusted H-index | | 0.281*** | 0.287*** |
| | | (0.05) | (0.05) |
| PhD supervisor in the committee | | | 0.369** |
| | | | (0.16) |
| PhD from the same institute | | | 0.311*** |
| | | | (0.10) |
| Gender × Discipline Fixed Effects | No | No | Yes |
| Gender × University Fixed Effects | No | Yes | Yes |
| Observations | 1357 | 1357 | 1357 |

$*p < 0.10, **p < 0.05, ***p < 0.01$. The gender variable is coded as 1 for a woman and 0 for a man. Robust standard errors in parentheses. The treatment variable is based on mean pre-reform gender parity at the field*university level. Variables interacted with the gender of the candidate should be interpreted as the differential effect of the variable for women with respect to men. Contests where only one candidate was ranked, or where only applicants of a specific gender applied are dropped from the analysis.

Table 3: Counterfactual Recruitments

| | True number | Simulated number | No reform for treated | No reform for treated and control |
|---|---|---|---|---|
| Man hired | 118 | 128 | 106 | 120 |
| Woman hired | 78 | 68 | 90 | 76 |
| Total | 196 | 196 | 196 | 196 |

These results show the gender of the true and simulated first-ranked candidates in the post-reform period under different scenarios. Column 1 represents the true numbers of men and women recruited, Column 2 the gender of the simulated first-ranked candidates from the estimates in column 3 of Table 2, with the 5 other measures of publications included. Column 3 represents the gender of the simulated first ranked candidates but imposes no effect of the reform for the treated group. Column 4 imposes no effect for both treatment and control group.

whether more women are hired when there are more women in the committee. I estimate the equation below at the contest level to see whether women are more or less likely to be first-ranked in the treatment group after the reform:

$$\mathrm{W}_{ij} = \alpha R_j T_j + \gamma T_j + \delta R_j + \beta_1 X_i + \epsilon_{ij}$$

where $W_{ij}$ is equal to 1 if the first ranked candidate is a woman, $X_i$ are control variables such as discipline fixed effects and or the share of female candidates, $\gamma$ represents the effect of being in the treatment group before the reform, $\delta$ the effect of the reform on the control group, and $\alpha$ is the parameter of interest, the effect of being in the treatment group after the reform relative to the control group. This equation is estimated through probit.

The results are presented in columns 1 and 2 of Table 4. I drop committees where only men or women applied in column 2 and add Discipline fixed effects. To partially control for differences in publications in this setting, I add a dummy variable that is equal to 1 if the candidate with the highest h-index is a woman. Though this is a sub-optimal way to control for publications, this variable has a large and significant effect on the probability that a woman is hired. The gender of the committee president is insignificant in all columns, as in the estimations at the candidate level. Our variable of interest, the effect of being in the treated group post-reform, has a significant negative effect on the probability of a woman being hired, just as in the rank-ordered logit. The point estimate is comparable in size to having 20% more female candidates.

This methodology might not be the most appropriate however, since I cannot control for individual specific characteristics, such as connections, in this specification. A solution to this problem is to estimate the probability of being hired through a conditional logit. A conditional logit estimates the probability that candidate $i$ is picked out of a pool of applicants. The specification estimated is the following:

$$\mathrm{F}_{ij} = \alpha R_j T_j g_i + \gamma T_j g_i + \delta R_j g_i + \beta_1 X_i + \beta_2 g_i Y_j + \epsilon_{ij}$$

where $F_{ij}$ is a variable equal to 1 if the candidate is ranked first, and 0 if not, $\gamma$ represents the effect of being in the treatment group before the reform, $\delta$ the effect of the reform on the control group, and $\alpha$ is the parameter of interest, the effect of being in

Table 4: Effect of the reform on the gender of the first-ranked candidate

| Dependent variable: | First-ranked candidate is female | | Ranked first | |
|---|---|---|---|---|
| Estimator: | Probit (Marginal) | | C.Logit | |
| | (1) | (2) | (3) | (4) |
| *Interacted with gender of the candidate for conditional logit* | | | | |
| **Post Reform × Treatment Group** | -0.171** | -0.207** | -1.677*** | -1.505** |
| | (0.08) | (0.10) | (0.61) | (0.66) |
| Post-Reform | 0.088* | 0.129** | 0.658* | 0.732* |
| | (0.05) | (0.06) | (0.37) | (0.39) |
| Treatment Group | 0.118* | 0.166 | 0.965* | 1.027 |
| | (0.07) | (0.14) | (0.51) | (0.79) |
| Female president | -0.025 | -0.058 | -0.561* | -0.638* |
| | (0.04) | (0.05) | (0.32) | (0.34) |
| Share of female candidates | 0.703*** | 0.626*** | 0.341 | 0.475 |
| | (0.08) | (0.16) | (0.76) | (0.83) |
| *At candidate level for c. logit* | | | | |
| H-index | 0.246*** | 0.300*** | 0.347*** | 0.360*** |
| | (0.03) | (0.04) | (0.07) | (0.08) |
| Controls | Professorship | Professorship | Yes | Yes |
| Discipline Fixed Effects | No | Yes | No | Yes |
| University Fixed Effects | Yes | Yes | Yes | Yes |
| Observations | 455 | 359 | 1357 | 1357 |

$*p < 0.10, **p < 0.05, ***p < 0.01$. Robust standard errors in parentheses. The gender variable is coded as 1 for a woman and 0 for a man. Variables interacted with the gender of the candidate should be interpreted as the differential effect of the variable for women with respect to men. Contests with applicants from a specific gender only are dropped in column 2. H-index variable for columns 1 and 2 is a dummy variable equal to 1 if the maximum h-index of all candidates belongs to a woman. It is the standardised age-adjusted h-index for columns 3 and 4. Controls in columns 3 and 4 include dummies for an application at the full professor level, for whether the candidate has a diploma from the university and for whether the candidates director is in the committee. The fixed effects are interacted with gender of the candidate for columns 3 and 4.

the treatment group after the reform relative to the control group. $g_i$ is a binary variable equal to 1 if the candidate is a woman, $X_i$ is a vector of individual characteristics of the candidate (such as the candidate's h-index.[30] or whether the candidate received his PhD from the university he is applying to), and $Y_j$ a vector of committee characteristics (such as the share of female applicants) that are interacted with $g_i$. The model is then estimated using a conditional logit.

The results are presented in columns 3 and 4 of Table 4. We can see that the reform has a large, significant and negative effect on the hiring of women. This effect is robust to the inclusion of other covariates and field fixed effects. Connections and publications have a positive effect on hiring, as expected. The results can be compared to the size of the coefficient on the h-index below. The h-index is standardised by field and seniority since the average h-index varies greatly along these dimensions. The negative effect of the reform is comparable to a 4 sd deviation drop in the h-index of candidates. However, this seems to be partially driven by the large positive effect of the reform on the control group.[31]

What do the preceding results imply for the effect size of increasing women in recruitment committees? Using dyads, as presented in Section D of the Appendix, I find that increasing the share of women in recruitment committees by 10% has the same negative effect as increasing the h-index difference between candidates by 2 standard deviations.

## 4.2 Effect of the reform on the candidate pool

A potential effect of the reform is to encourage more women to apply for academic positions, if they believe that more feminised committees will be better inclined towards them. However the results above could be biased if the composition of the applicant pool changes drastically with the reform. To check whether the approach used to estimate the effect of the reform is valid, I test whether the gender composition of the applicants changes significantly before and after the reform in both groups. There are in fact changes in the

[30] I use the age-adjusted h-index of candidates since this is the publication measure that is the most significant when including all other measures, as seen in Table H.20 in the Appendix.

[31] This is slightly worrying, but also consistent with the IV estimates in the Appendix that show the control group, when analysed at the candidate level, has a reduction in the share of female jurors in the first stage. The rank-ordered logit estimates, which use the data more efficiently, show no significant effect in the control group.

Table 5: Effect of the reform on the candidate pool

| Dependent variable: | Gender composition of candidates | | | | | |
|---|---|---|---|---|---|---|
| Years | 2009-2018 | | | 2013-2016 | | |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Post-Reform | 0.014 | 0.068* | 0.021 | 0.033 | 0.003 | 0.038 |
| | (0.03) | (0.04) | (0.03) | (0.03) | (0.04) | (0.03) |
| Female president | 0.051 | -0.014 | 0.019 | 0.072* | -0.050 | 0.025 |
| | (0.04) | (0.04) | (0.03) | (0.04) | (0.04) | (0.03) |
| Professorship | -0.157*** | -0.126*** | -0.145*** | -0.175*** | -0.100*** | -0.144*** |
| | (0.04) | (0.03) | (0.02) | (0.04) | (0.03) | (0.03) |
| Treatment Group | | | -0.032 | | | 0.021 |
| | | | (0.04) | | | (0.06) |
| Post-Reform × Treatment Group | | | 0.040 | | | -0.041 |
| | | | (0.05) | | | (0.05) |
| Group | Control Group | Treatment Group | Both | Control Group | Treatment Group | Both |
| Gender × Discipline Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes |
| Gender × University Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 235 | 220 | 455 | 166 | 130 | 296 |

$*p < 0.10, **p < 0.05, ***p < 0.01$. The gender variable is coded as 1 for a woman and 0 for a man. Robust standard errors in parentheses. The treatment variable is based on mean pre-reform gender parity at the field*university level. The last 3 columns are restrained to 2013-2016.

gender composition of the applicant pool after the reform, though as shown in column 3, there are no significant differences between treated and control group in this respect. This could also be due to the long time span in my data, with committees from 2009 to 2018.

One way to investigate this problem is to restrict the time span to the period around the reform, from 2013 to 2017. In this time frame, there are no significant changes in the gender composition of the applicant pool after the reform in both treated and control groups, and as seen in column 6 the point estimate for the difference between treated and control group is in fact negative. This indicates that the changes in the applicant pool seen in columns 1 and 2 could be due to longer-run trends in the propensity of women to apply for these positions rather than the reaction of candidates to the reform. We can conclude that the reform doesn't seem to encourage more women to apply to professorial positions, at least in the very short-term.

Nevertheless, Table 6 presents results using the years immediately leading up to and following the reform to alleviate this concern. Columns 1 and 2 present the results with this limited time frame, only including the years from 2013 to 2016. The point estimates are once again very close to those presented in the other tables, so we can probably rule out the idea that the estimated effect of the reform is due either to longer-run trends, or due to the changing nature of the application pool as a reaction to the reform.

Table 6: Effect of the quota on the rank of female candidates - Limited years, top 3 candidates and continuous specification

| Dependent variable: | Ranked first | Rank | Ranked first | Rank | Ranked first | Rank |
|---|---|---|---|---|---|---|
| Estimator: | C. Logit | RO. Logit | C. Logit | RO. Logit | C. Logit | RO. Logit |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| *Interacted with gender of the candidate* | | | | | | |
| **Post Reform × Treatment Group** | -2.234*** | -0.906** | -2.216*** | -1.696*** | | |
| | (0.80) | (0.46) | (0.83) | (0.63) | | |
| Post-Reform | 0.819* | 0.181 | 0.913 | 0.254 | 0.667* | 0.069 |
| | (0.50) | (0.26) | (0.56) | (0.43) | (0.37) | (0.22) |
| Treatment Group | 1.427 | 1.633** | 0.447 | 1.174 | | |
| | (0.99) | (0.66) | (0.94) | (0.76) | | |
| **Distance to threshold-post reform** | | | | | -15.616** | -7.078* |
| | | | | | (6.86) | (3.95) |
| Distance to threshold-pre reform | | | | | 9.603 | 7.761 |
| | | | | | (7.85) | (5.33) |
| *Variables at the candidate level* | | | | | | |
| Standardised age-adjusted H-index | 0.457*** | 0.320*** | 0.386*** | 0.280*** | 0.358*** | 0.285*** |
| | (0.09) | (0.06) | (0.11) | (0.09) | (0.08) | (0.05) |
| Time | 2013-2016 | 2013-2016 | 2013-2016 | 2013-2016 | 2009-2018 | 2009-2018 |
| Ranks | All | All | Top 3 | Top 3 | All | All |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes |
| Gender × Discipline Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes |
| Gender × University Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 926 | 926 | 597 | 597 | 1357 | 1357 |

$*p < 0.10, **p < 0.05, ***p < 0.01$. The gender variable is coded as 1 for a woman and 0 for a man. Robust standard errors in parentheses. The treatment variable is based on mean pre-reform gender parity at the field*university level. Variables interacted with the gender of the candidate should be interpreted as the differential effect of the variable for women with respect to men. Contests where only one candidate was ranked, or where only applicants of a specific gender applied are dropped from the analysis. Only the two years either side of the reform are included in columns 1-4. Only the top 3 candidates are considered in column 3-4. Controls include the standardised, age-adjusted h-index of the candidates, and dummies for an application at the full professor level, for whether the candidate has a diploma from the university and for whether the candidates director is in the committee.

## 4.3 Top ranks only

One concern in using ranked data is that some committees may not care about how they rank lower quality candidates, since they are very unlikely to be hired. Our results could thus be driven by inattention from jury members towards candidates less likely to be recruited rather than discrimination.[32] To investigate this concern, I present results in columns 3 and 4 of Table 6 using again the conditional and rank-ordered logit specifications, but keeping only the top 3 candidates from each committee, and dropping candidates that were auditioned and not ranked. I keep the two years either side of the reform only as in Columns 1 and 2. In this specification, the reform still has a significant effect and the point estimates are even larger than those that consider the full ranking. We can therefore rule out that the effect of the reform is driven by inattention or lack of effort from juries when ranking candidates that are very unlikely to be hired.

## 4.4 Continuous specification

Another way of approaching the problem is to consider a continuous specification of the treatment variable. Although using a binary variable to determine assignment to treatment enables an easy interpretation of the treatment variable, fields that were far away from the cut-off before the reform are weighted in the same away as those that were close.[33] A standard way to estimate the effect of treatment intensity would be to take the difference in the outcome variable post- and pre-reform, and then regress this new variable on the distance to the threshold pre-reform.[34] It is impossible to do this using my data however since my observations are at the field level, and I therefore cannot difference my outcome variable unless I collapse the data at field or field university level. I use the following specification instead to estimate the effect of the reform on ranks of women:

[32] See Bartoš et al. (2016) for evidence that recruiters' attention decreases when considering minority job applicants.

[33] We might also be worried by the comparability of fields for which the mean share of women pre-reform is very far away from the threshold. I consider this possibility in Table H.25 in the Appendix, keeping only labs where the pre-reform share of women is between 35 and 45%, with similar results to those presented in the main part of the article.

[34] See Casas-Arce and Saiz (2015) for an example.

$$\text{DV}_{ij} = \xi g_i \max\{0, 0.4 - \text{jury}_x^{pre-2015}\} + \mu_{ij} g_i^{post-2015} \max\{0, 0.4 - \text{jury}_x^{pre-2015}\}$$

$$+\beta_1 X_i + \beta_2 g_i Y_j + \epsilon_{ij}$$

where $\xi$ captures the effect of the treatment intensity before the reform and jury is the mean share of female jurors in a field before 2015, and DV is the dependent variable of interest, the rank or whether the candidate finished first depending on the specification. Our parameter of interest here is $\mu_{ij}$, which should capture the post-reform effect of the treatment. The results for the conditional logit and the rank-ordered logit can be seen in columns 5 and 6 of Table 6, and are similar to those presented in the previous tables, though the results for the ranks here are significant at the 10 percent level only. An interpretation of the results presented here is that being in a field that was 10 percentage points away from the cut-off before the reform has an effect similar to a 3 unit decrease in age-adjusted h-index. The results remain quantitatively similar when using distance to the threshold on the other side rather than normalising all labs with above 40% women in committees before the reform to 0.

## 4.5 Other dimensions of the reform

In this section, I analyse the effect of the reform on other dimensions than the recruitment of women. A first question is related to how committees are formed as a reaction to the quota. Are the new jury members coming from inside or outside the department that is recruiting, taking into account the fact that at least half the members of the committees have to be from outside the department? In Table 7, I regress the share of women among internal members and external members on the treatment group and reform variables, in columns 1 and 2 respectively. In both cases, the effect of being affected by reform is positive, though the point estimate for internal members is larger.[35] We can conclude that as a reaction to the reform, committees increase both the number of external and internal

[35]Note that since the IV estimate is essentially a Wald estimator, using the IV strategy from above to instrument for the internal and external share would necessarily give a larger effect size for the group that was least affected by the reform.

female committee members, though they seem to draw more from their colleagues inside the department. This result implies that women inside the departments most affected by the reform have more administrative work post-2015 relative to their male colleagues, which could harm their ability to publish.

Does the reform affect the h-index of committee members? We could think that the new female jurors have less publications than the male professors that they replace. In column 3 of Table 7, I regress the mean h-index of jurors by committees on the treatment and reform dummies. There doesn't appear to be much of an effect of the reform, with the negative treatment effect on the mean h-index of jurors both small (around 0.1 of a standard deviation in the h-index) and insignificant. Combined with the results on internal versus external members, this shows that two concerns we could have about the reform; that jury president choose junior women on purpose to negate the effects of the reform, and that the new female jurors are less qualified than the members they replace, degrading the overall "quality" of the committee, do not seem to hold in the data.

We might also be worried that the way committees are formed changes with the reform. In columns 4 and 7 of Table 7, I try to see whether the reform affected the number of members in committees and the likelihood for the president to be female, using the same diff-in-diff strategy as in the main part of the article in columns 1 and 2. For both variables, there is no significant effect of the reform. The lack of change in female presidents is consistent with the literature on boardroom quotas, which points out the lack of women in the positions of most power. Columns 5 and 6 deal with the issue of endogeneity in the choice of the jury president. The way they are chosen could have changed with the reform, which could bias our results since it is included in most regressions. However, there is no detectable effect of the reform on the age or the h-index of committee presidents, which along with the results in column 7 lends no support to this hypothesis.

## 4.6 Differentiated effects of networks

Qualitative research has shown that women may benefit less from academic connections than men do, and that thesis directors are more likely to abstain during committee dis-

Table 7: Increase in the share of internal and external members, effect on quality of jurors and selection of presidents and committee size

| Dependent variable: | Share of women among internal members | Share of women among external members | Mean standardised h-index of jurors | Number of committee members | Age of jury president | H-index of jury president | Gender of jury president |
|---|---|---|---|---|---|---|---|
| Estimator: | | | OLS | | | | Probit (marginal effects) |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
| **Post Reform × Treatment Group** | 0.174*** | 0.110*** | -0.110 | 0.595 | 1.130 | -0.019 | -0.055 |
| | (0.03) | (0.04) | (0.10) | (0.42) | (1.52) | (0.22) | (0.08) |
| Treatment Group | -0.167*** | -0.190*** | 0.238* | -0.341 | 1.083 | -0.119 | 0.048 |
| | (0.03) | (0.03) | (0.14) | (0.61) | (1.96) | (0.29) | (0.12) |
| Post-Reform | -0.006 | 0.006 | -0.034 | -0.281 | 0.053 | -0.191 | 0.018 |
| | (0.02) | (0.03) | (0.06) | (0.28) | (0.93) | (0.14) | (0.05) |
| Baseline | 0.520*** | 0.437*** | | | | | |
| | (0.02) | (0.02) | | | | | |
| Discipline Fixed Effect | No | No | No | Yes | Yes | Yes | Yes |
| University Fixed Effect | No | No | Yes | Yes | Yes | Yes | Yes |
| Observations | 455 | 455 | 455 | 455 | 447 | 455 | 455 |

$*p < 0.10, **p < 0.05, ***p < 0.01$. The gender variable is coded as 1 for a woman and 0 for a man. Robust standard errors in parentheses. The treatment variable is based on mean pre-reform gender parity at the field*university level. H-index standardised by discipline and seniority. Robust standard errors in parentheses. Last 4 columns correct for professorship.

Table 8: Gendered effects of networks

| Dependent variable: | Rank | | | |
|---|---|---|---|---|
| Estimator: | R.O. Logit | | | |
| | (1) | (2) | (3) | (4) |
| *Variable at the candidate level* | | | | |
| Standardised age-adjusted H-index | 0.274*** | 0.279*** | 0.279*** | 0.283*** |
| | (0.05) | (0.05) | (0.05) | (0.05) |
| Candidate is a woman | -0.191** | -0.132 | -0.216** | -0.145 |
| | (0.08) | (0.22) | (0.09) | (0.22) |
| PhD supervisor in the committee | 0.736*** | 0.758*** | | |
| | (0.24) | (0.25) | | |
| PhD supervisor in the committee | -0.517* | -0.564* | | |
| × candidate is female | (0.30) | (0.31) | | |
| PhD from the same institute | | | 0.356** | 0.376*** |
| | | | (0.14) | (0.14) |
| PhD from the same institute | | | -0.060 | -0.091 |
| × candidate is female | | | (0.19) | (0.20) |
| Gender × Discipline Fixed Effects | No | Yes | No | Yes |
| Observations | 1357 | 1357 | 1357 | 1357 |

$*p < 0.10, **p < 0.05, ***p < 0.01$. The gender variable is coded as 1 for a woman and 0 for a man. Robust standard errors in parentheses. The treatment variable is based on mean pre-reform gender parity at the field*university level. Contests where only one candidate was ranked, or where only applicants of a specific gender applied are dropped from the analysis.

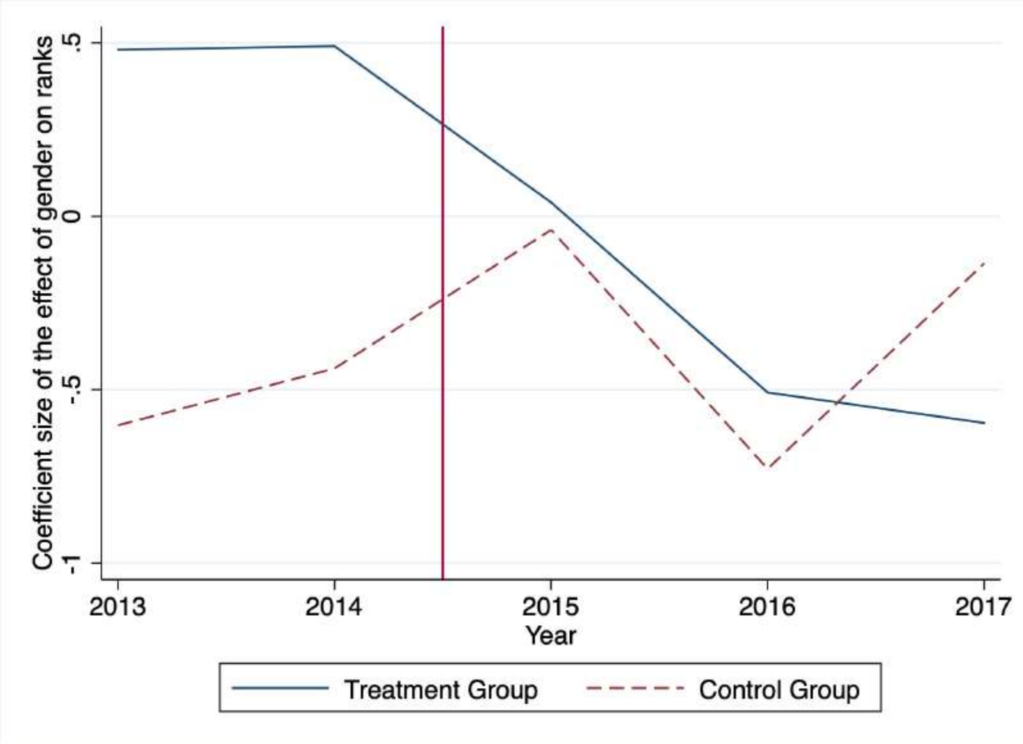
cussions if their PhD student is a woman.[36] In my dataset, women are more likely to have their director in the committee than men do,[37] but the gains from this connection may not be the same for both genders. I directly test this hypothesis in Table 8 by interacting the effect of having one's advisor in the committee with the gender of the candidate, using the rank-ordered logit framework. The results are striking: women benefit much less from having their supervisor in the committee than men do. In fact, the point estimates imply that the positive effect of having one's thesis director in the committee for a woman is almost exactly cancelled out by the negative effect of being a woman. In other words, having one's director in the committee enables women to be ranked the same as men, though the effect is significantly different from the one for men only at the 10 percent level. There are no significant or large differences for the effect of having received a PhD from the institution where one is applying.

# 5 Discussion - How should we interpret this result?

I provide a graphical interpretation of the main results in Figure 3. In this figure, I plot the coefficients from a rank-ordered logit of being a woman, in the two years before and

[36] See Sautier (2019), whose research on this topic is discussed in Deschamps et al. (2020).
[37] 48 women have their directors in the committee against 42 men.

Figure 3: The effect of gender on ranks from a rank-ordered logit, controlling for publications



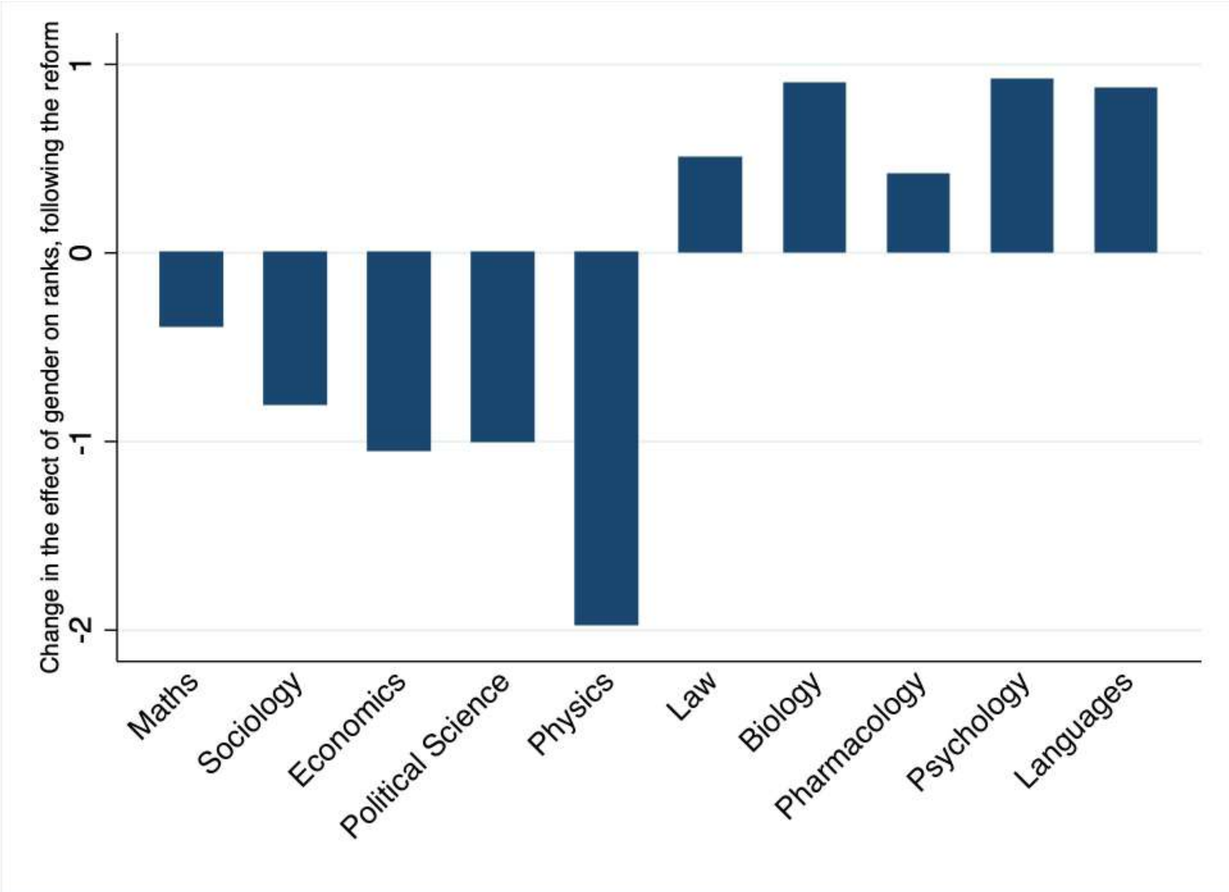

after the reform, controlling for the h-index of candidates. The treatment group consists of recruitments in disciplines that are affected by the reform, while the control group is made up of disciplines that were already respecting gender parity in committees, and should therefore not be affected by the reform, as explained above. Slightly surprising is that pre-reform, fields that were not feminised actually tended to favour women. This result is similar to those reported by Breda and Hillion (2016), which showed that women were favoured in oral exams for posts as secondary-school teachers in disciplines that had few women instructors. In disciplines affected by the reform, the rank of women decreases, whereas women are ranked slightly higher in the control group.

This effect does not seem to be driven by outliers in fields. In Figure 4, I look at the effect of the reform on the ranks of women by discipline,[38] with disciplines sorted from left to right according to their pre-reform average share of female jurors. The figure bears out what we see in the regression: disciplines that had few women jurors, and thus were more affected by the quota seem to be less favourable to women after the reform, while

[38]For clarity, fields with a small number of observations have been dropped from this figure. This figure plots the change in the effect of gender on ranks using the narrowed time-frame, 2013-2016, controlling for the h-index of professors and the gender of the jury president.

disciplines that were already respecting the quota (such as Psychology and Languages), seem to have been positively affected by the reform.

Figure 4: Effect of the reform on the ranks of women by discipline. Disciplines most affected by the reform are ordered from left to right.




One issue in interpreting these results is that we cannot identify what is driving these worse outcomes, especially since most of the jury level characteristics seem to have little to no effect. One possibility is that women in positions of authority have opposite-gender preferences, a phenomenon that has already been analysed in the sociology literature under the expression "Queen Bee" syndrome.[39] An alternative to the Queen Bee syndrome is that women are penalised when they decide to promote women, as discussed in Hekman et al. (2016). The effect that we see is then due to women internalising the retribution they could face from male colleagues if they promote women, rather than opposite-gender preferences. Akerlof and Kranton (2000) postulate instead that men discriminate against women when their identities are threatened. It is possible that this is the case in our setting, since having more women in what were traditionally masculine settings could be

[39]see e.g Staines et al. (1974).

seen as a threat to the masculinity of the jurors. The interpretation of our effects in this case would not be of opposite-gender bias in recruiting, but of own-gender bias from men that appears only when the gender identities of jurors are threatened. Our results are consistent with all three hypotheses.

However, an element that leans towards the third hypothesis is presented in Table 9. In these tables I estimate the following specification:

$$\mathrm{DV}_{ij} = \theta R_j T_j g_i g_j + \alpha R_j T_j g_i + \gamma T_j g_i + \lambda T_j g_i g_j + \delta R_j g_i + \kappa R_j g_i g_j + \beta_1 X_i + \beta_2 g_i Y_j + \epsilon_{ij}$$

where $g_j$ is a dummy variable equal to 1 if the jury president is female. Our parameters of interest here are $\alpha$ and $\theta$, the effect of the reform depending on the gender of the jury president. There is a significant difference between the effect of the reform for committees with male presidents and those with female presidents in the treatment group. The negative effect of the reform is entirely driven by committees with male presidents, for both the conditional and rank-ordered logit specifications. Recall that Table 7 shows no differences in the likelihood of women to be presidents, or significant changes in the age or h-indices of jury presidents, so it is unlikely that these results are driven by changes in selection of jury presidents following the reform. Results using jury fixed effects are similar and are presented in Section C of the Appendix. This result is consistent with the idea that men change their behaviour as a result of the reform.[40]

Irrespective of the precise channel that causes this effect and consistently with the previous literature, the results presented in this article strongly suggest that gender quotas and other coercive measures may have unintended effects. In this specific case, increasing the proportion of female jurors has led to a negative effect on the ranking of women by committees.

[40] We can also revisit the effect of the reform on first-ranked candidates in the probit specification. Interacting the effect of the reform with the gender of the jury president leads to similar results on hiring. These results are presented in Table H.19 of Appendix H.

Table 9: Effect of the reform on women by gender of the jury president

| Dependent variable: | Ranked first | | Rank | |
|---|---|---|---|---|
| Estimator: | C. Logit | | R.O. Logit | |
| | (1) | (2) | (3) | (4) |
| *Interacted with gender of the candidate* | | | | |
| **Post-Reform × Treatment Group** | -2.430*** | -2.715*** | -1.277*** | -1.190** |
| | (0.70) | (0.77) | (0.42) | (0.47) |
| **Post-Reform × Treatment Group** | 3.292*** | 3.979*** | 1.296** | 1.287* |
| **× Female President** | (1.28) | (1.35) | (0.63) | (0.69) |
| Treatment Group × Female president | -2.077** | -2.489** | -0.799* | -0.770 |
| | (0.97) | (1.04) | (0.43) | (0.49) |
| Treatment Group | 1.321** | 1.962** | 0.840** | 1.739*** |
| | (0.57) | (0.86) | (0.35) | (0.62) |
| Control Group × Post-Reform | 1.145** | 1.379*** | 0.387 | 0.351 |
| | (0.46) | (0.51) | (0.25) | (0.29) |
| Control Group × Post-Reform | -1.435* | -1.707** | -0.539 | -0.534 |
| × Female president | (0.75) | (0.87) | (0.43) | (0.49) |
| Share of female candidates | -0.262 | 0.803 | -0.255 | -0.241 |
| | (0.76) | (0.84) | (0.40) | (0.43) |
| Professorship | | 0.957*** | | 0.327* |
| | | (0.32) | | (0.19) |
| *Variable at the candidate level* | | | | |
| Candidate is a woman | -0.753 | -1.463 | -0.336 | -0.212 |
| | (0.58) | (0.95) | (0.29) | (0.47) |
| Candidate is a woman × Female president | 0.431 | 0.418 | 0.296 | 0.198 |
| | (0.58) | (0.65) | (0.30) | (0.35) |
| Standardised age-adjusted h-index | | 0.376*** | | 0.286*** |
| | | (0.08) | | (0.05) |
| PhD supervisor in the committee | | 0.435 | | 0.356** |
| | | (0.27) | | (0.16) |
| PhD from the same institute | | 0.590*** | | 0.312*** |
| | | (0.18) | | (0.10) |
| Gender × Discipline Fixed Effects | No | Yes | No | Yes |
| Gender × University Fixed Effects | No | Yes | No | Yes |
| Observations | 1357 | 1357 | 1357 | 1357 |

$*p < 0.10, **p < 0.05, ***p < 0.01$. The gender variable is coded as 1 for a woman and 0 for a man. Robust standard errors in parentheses. The treatment variable is based on mean pre-reform gender parity at the field*university level. Variables interacted with the gender of the candidate should be interpreted as the differential effect of the variable for women with respect to men. Contests where only one candidate was ranked, or where only applicants of a specific gender applied are dropped from the analysis.

# 6 Conclusion

In this article, I have showed that imposing gender quotas in recruitment committees does not help the recruitment of women, and may actually harm their careers. I find suggestive evidence that men vote against women as a reaction to the quota. The quota does not seems to affect the average quality of the jury, the likelihood of women to apply for academic positions or to be jury president.

The results do not preclude there being discrimination against women in hiring,[41] but they do suggest that there may not be an easy solution to the problem of under-representation of women in academia.

# References


ABREVAYA, J. AND D. S. HAMERMESH (2012): "Charity and Favoritism in the Field: Are Female Economists Nicer (To Each Other)?" *The Review of Economics and Statistics*, 94, 202–207.

AKERLOF, G. A. AND R. E. KRANTON (2000): "Economics and identity," *The Quarterly Journal of Economics*, 115, 715–753.

AZMAT, G. (2014): "Gender diversity in teams," *IZA World of Labor*.

BAGUES, M. AND P. CAMPA (2017): "Can gender quotas in candidate lists empower women? Evidence from a regression discontinuity design," .

BAGUES, M., M. SYLOS-LABINI, AND N. ZINOVYEVA (2017): "Does the gender composition of scientific committees matter?" *American Economic Review*, 107, 1207–38.

BAGUES, M. F. AND B. ESTEVE-VOLART (2010): "Can Gender Parity Break the Glass Ceiling? Evidence from a Repeated Randomized Experiment," *Review of Economic Studies*, 77, 1301–1328.

BAGUES, M. F. AND N. ZINOVYEVA (2015): "The Role of Connections in Academic Promotions," *American Economic Journal: Applied Economics*, 124(5).

BARTOŠ, V., M. BAUER, J. CHYTILOVÁ, AND F. MATĚJKA (2016): "Attention discrimination: Theory and field experiments with monitoring information acquisition," *American Economic Review*, 106, 1437–75.

BEAMAN, L., R. CHATTOPADHYAY, E. DUFLO, R. PANDE, AND P. TOPALOVA (2009): "Powerful women: does exposure reduce bias?" *The Quarterly Journal of Economics*, 124, 1497–1540.

BEAMAN, L., E. DUFLO, R. PANDE, AND P. TOPALOVA (2012): "Female leadership raises aspirations and educational attainment for girls: A policy experiment in India," *Science*, 335, 582–586.

BEGGS, S., S. CARDELL, AND J. HAUSMAN (1981): "Assessing the potential demand for electric cars," *Journal of Econometrics*, 17, 1–19.

BERTRAND, M., S. E. BLACK, S. JENSEN, AND A. LLERAS-MUNEY (2018): "Breaking the glass ceiling? The effect of board quotas on female labor market outcomes in Norway," *Review of Economic Studies*.


[41] It seems in fact quite likely that this is the case. See Table B.6 for instance, where the coefficient in column 1 for the effect of being a woman on rank is negative, though that result also includes the negative effect of the reform.

Blau, F. D. and L. M. Kahn (2017): "The gender wage gap: Extent, trends, and explanations," *Journal of Economic Literature*, 55, 789–865.

Boring, A. (2017): "Gender biases in student evaluations of teaching," *Journal of Public Economics*, 145, 27–41.

Bosquet, C., P.-P. Combes, and C. Garcia-Peñalosa (2018): "Gender and Promotions: Evidence from Academic Economists in France," *Scandinavian Journal of Economics*.

Breda, T. and M. Hillion (2016): "Teaching accreditation exams reveal grading biases favor women in male-dominated disciplines in France," *Science*, 353, 474–478.

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

Card, D., S. DellaVigna, P. Funk, and N. Iriberri (2019): "Are referees and editors in economics gender neutral?" *The Quarterly Journal of Economics*.

Casas-Arce, P. and A. Saiz (2015): "Women and power: unpopular, unwilling, or held back?" *Journal of Political Economy*, 123, 641–669.

Chenu, A. and O. Martin (2016): "Le plafond de verre chez les enseignants-chercheurs en sociologie et démographie," *Travail, genre et sociétés*, 133–156.

Colussi, T. (2018): "Social Ties in Academia: A Friend Is a Treasure," *The Review of Economics and Statistics*, 100, 45–50.

Combes, P.-P., L. Linnemer, and M. Visser (2008): "Publish or peer-rich? The role of skills and networks in hiring economics professors," *Labour Economics*, 15, 423–441.

Dalvit, N., A. Patel, and J. Tan (2018): "Intra-firm Hierarchies and Gender Gaps," .

Deschamps, P., A. Revillard, M. Sautier, and M. Arbogast (2020): "Les quotas de genre dans le monde académique," *Débats du LIEPP N.5*.

Duflo, E. (2000): "Schooling and labor market consequences of school construction in Indonesia: Evidence from an unusual policy experiment," Tech. rep., National Bureau of Economic Research.

Edo, A., N. Jacquemet, and C. Yannelis (2015): "Language skills and homophilous hiring discrimination: Evidence from gender and racially differentiated applications," *Review of Economics of the Household*, 1–28.

Ellemers, N., H. Van den Heuvel, D. de Gilder, A. Maass, and A. Bonvini (2004): "The underrepresentation of women in science: Differential commitment or the queen bee syndrome?" *British Journal of Social Psychology*, 43, 315–338.

Eyméoud, J.-B. and P. Vertier (2017): "Gender Biases: Evidence from a Natural Experiment in French Local Elections," *LIEPP Working Paper*.

Feld, J., N. Salamanca, and D. S. Hamermesh (2016): "Endophilia or exophobia: beyond discrimination," *The Economic Journal*, 126, 1503–1527.

Geuna, A. and S. Shibayama (2015): "Moving out of academic research: why do scientists stop doing research?" in *Global mobility of research scientists*, Elsevier, 271–303.

Goldin, C. (2014): "A grand gender convergence: Its last chapter," *American Economic Review*, 104, 1091–1119.

Hausman, J. A. and P. A. Ruud (1987): "Specifying and testing econometric models for rank-ordered data," *Journal of Econometrics*, 34, 83–104.

Hekman, D. R., S. Johnson, M. Der Foo, and W. Yang (2016): "Does diversity-valuing behavior result in diminished performance ratings for nonwhite and female leaders?" *Academy of Management Journal*, amj–2014.

Hsieh, C.-T., E. Hurst, C. I. Jones, and P. J. Klenow (2019): "The allocation of talent and US economic growth," *Econometrica*, 87, 1439–1474.

Ianni, G. (2010): "Scholar H-index Calculator," .

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

Lincoln, A. E., S. Pincus, J. B. Koster, and P. S. Leboy (2012): "The Matilda Effect in science: Awards and prizes in the US, 1990s and 2000s," *Social Studies of Science*, 42, 307–320.

Maida, A. and A. Weber (2019): "Female Leadership and Gender Gap within Firms: Evidence from an Italian Board Reform," .

Niederle, M. and L. Vesterlund (2007): "Do Women Shy Away From Competition? Do Men Compete Too Much?" *The Quarterly Journal of Economics*, 122, 1067–1101.

Paola, M. D. and V. Scoppa (2015): "Gender Discrimination and Evaluators' Gender: Evidence from Italian Academia," *Economica*, 82, 162–188.

Rebérioux, A. and G. Roudaut (2017): "Gender Quota and Inequalities inside the Boardroom," .

Sautier, M. (2019): "Hiring beyond borders ? The Moral Economies of Mobility in Academia in Comparative Perspective," *Thesis in Progress*.

Staiger, D. and J. Stock (1997): "Instrumental Variables Regression with Weak Instruments," *Econometrica*, vol.65, 557–586.

Staines, G., C. Tavris, and T. E. Jayaratne (1974): "The queen bee syndrome," *Psychology Today*, 7, 55–60.

Thurstone, L. L. (1927): "A Law of Comparative Judgement," *Psychological Review*, 34, 278–286.

Williams, W. M. and S. 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, 5360–5365.

Zeltzer, D. (2020): "Gender Homophily in Referral Networks: Consequences for the Medicare Physician Pay Gap," *American Economic Journal: Applied Economics*.

# A Descriptive statistics

In this section I present additional descriptive statistics. Table A.1 shows descriptive statistics on applications and ranks. Committees with missing application data are not included in this table. As can be seen from column 2, excluding these committees changes little to the share of women hired Table A.2 shows some statistics on the characteristics of candidates, such as the average h-index by field, or the number of candidates that have a supervisor in the committee. Table A.3 shows how the share of women in recruitment committees and as jury presidents evolves with the reform. Table A.4 shows the number of observations per field and year. Although I have some early observations from 2009-2012, the bulk of the data is from post-2013.

Table A.1: Descriptive statistics on applications and hiring

| | Share of female applicants | Share of women hired | Share of women ranked |
|---|---|---|---|
| Biology | 0.54 | 0.55 | 0.57 |
| Business | 0.42 | 0.27 | 0.55 |
| Chemistry | 0.39 | 0.46 | 0.47 |
| Economics | 0.26 | 0.16 | 0.25 |
| Education | 0.63 | 0.54 | 0.65 |
| Engineering | 0.23 | 0.07 | 0.20 |
| History | 0.28 | 0.36 | 0.31 |
| Languages | 0.56 | 0.47 | 0.63 |
| Law | 0.46 | 0.42 | 0.44 |
| Maths | 0.19 | 0.15 | 0.21 |
| Pharmacology | 0.53 | 0.54 | 0.64 |
| Physical Education | 0.15 | 0.11 | 0.15 |
| Physics | 0.20 | 0.23 | 0.18 |
| Political Science | 0.35 | 0.24 | 0.43 |
| Psychology | 0.62 | 0.61 | 0.66 |
| Sociology | 0.43 | 0.46 | 0.50 |
| Total | 0.40 | 0.37 | 0.44 |
| Missing Candidates included? | No | No | No |

Percentages represent averages over all committees in a field.

Table A.2: Descriptive statistics on academic connections

| | # Total Candidates | Candidates w/ supervisor in jury | Candidates w/ PhD from the institution | Mean h-index |
|---|---|---|---|---|
| Biology | 204 | 19 | 79 | 8.84 |
| Business | 52 | 1 | 5 | 2.94 |
| Chemistry | 50 | 2 | 18 | 14.4 |
| Economics | 76 | 3 | 14 | 8.22 |
| Education | 53 | 2 | 11 | 9.98 |
| Engineering | 59 | 2 | 14 | 8.05 |
| History | 41 | 0 | 7 | 6.51 |
| Languages | 148 | 17 | 27 | 2.71 |
| Law | 102 | 8 | 34 | 2.73 |
| Maths | 267 | 7 | 39 | 6.51 |
| Pharmacology | 95 | 5 | 13 | 8.19 |
| Physical Education | 36 | 1 | 9 | 5.78 |
| Physics | 107 | 7 | 36 | 9.98 |
| Political Science | 64 | 3 | 14 | 9.27 |
| Psychology | 82 | 10 | 34 | 5.48 |
| Sociology | 112 | 5 | 11 | 9.03 |
| Total | 1548 | 90 | 365 | 6.95 |

Data on candidates' PhDs is scraped from online sources.

Table A.3: Effect of the reform on the mean share of women jurors and number of female presidents

| | Mean share of female jurors | | Share of female presidents | |
|---|---|---|---|---|
| Discipline | Pre-reform | Post-reform | Pre-reform | Post-reform |
| Biology | 0.47 | 0.49 | 0.32 | 0.43 |
| Business | 0.39 | 0.52 | 0.00 | 0.29 |
| Chemistry | 0.37 | 0.46 | 0.80 | 0.31 |
| Economics | 0.32 | 0.47 | 0.00 | 0.12 |
| Education | 0.37 | 0.49 | 0.25 | 0.14 |
| Engineering | 0.33 | 0.43 | 0.00 | 0.15 |
| History | 0.35 | 0.45 | 0.40 | 0.50 |
| Languages | 0.55 | 0.50 | 0.56 | 0.42 |
| Law | 0.38 | 0.47 | 0.14 | 0.05 |
| Maths | 0.30 | 0.41 | 0.31 | 0.29 |
| Pharmacology | 0.46 | 0.49 | 0.44 | 0.38 |
| Physical Education | 0.45 | 0.47 | 0.14 | 0.00 |
| Physics | 0.34 | 0.44 | 0.08 | 0.19 |
| Political Science | 0.31 | 0.50 | 0.33 | 0.29 |
| Psychology | 0.52 | 0.49 | 0.21 | 0.62 |
| Sociology | 0.32 | 0.50 | 0.40 | 0.38 |
| Total | 0.40 | 0.47 | 0.30 | 0.30 |

Percentages represent averages over all committees in a field.

Table A.4: Observations by Field and Year

| Year | 2009 | 2010 | 2011 | 2012 | 2013 | 2014 | 2015 | 2016 | 2017 | 2018 | Total |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Biology | 0 | 0 | 0 | 5 | 12 | 11 | 12 | 11 | 5 | 9 | 65 |
| Business | 0 | 0 | 0 | 1 | 0 | 3 | 4 | 2 | 4 | 4 | 18 |
| Chemistry | 0 | 0 | 0 | 2 | 0 | 3 | 6 | 3 | 2 | 2 | 18 |
| Economics | 2 | 1 | 0 | 2 | 1 | 5 | 5 | 6 | 4 | 1 | 27 |
| Education | 0 | 0 | 0 | 0 | 0 | 4 | 3 | 4 | 5 | 2 | 18 |
| Engineering | 0 | 0 | 0 | 0 | 0 | 5 | 4 | 2 | 2 | 5 | 18 |
| History | 1 | 2 | 1 | 3 | 1 | 2 | 1 | 0 | 1 | 0 | 12 |
| Languages | 0 | 0 | 0 | 0 | 7 | 9 | 9 | 10 | 5 | 0 | 40 |
| Law | 0 | 3 | 1 | 1 | 1 | 8 | 3 | 5 | 6 | 5 | 33 |
| Maths | 0 | 0 | 0 | 1 | 7 | 8 | 16 | 9 | 7 | 9 | 57 |
| Pharmacology | 0 | 0 | 0 | 5 | 1 | 3 | 5 | 6 | 5 | 5 | 30 |
| Physical Education | 0 | 0 | 0 | 1 | 2 | 4 | 0 | 1 | 2 | 3 | 13 |
| Physics | 0 | 0 | 0 | 0 | 6 | 6 | 5 | 6 | 1 | 4 | 28 |
| Political Science | 1 | 3 | 1 | 3 | 2 | 2 | 1 | 2 | 3 | 1 | 19 |
| Psychology | 0 | 0 | 0 | 2 | 5 | 7 | 6 | 6 | 3 | 1 | 30 |
| Sociology | 0 | 2 | 1 | 5 | 6 | 7 | 0 | 5 | 2 | 1 | 29 |
| Total | 4 | 11 | 4 | 31 | 51 | 87 | 80 | 78 | 57 | 52 | 455 |

Table B.5: Correlation between the gender of jurors and the probability of women being first-ranked

| Dependent variable: | Gender(W) of the first ranked candidate | | | |
|---|---|---|---|---|
| Estimator: | Probit | Probit | Probit | Probit |
| Marginal effects | (1) | (2) | (3) | (4) |
| Share of women in the committee | 0.787*** | 0.006 | 0.098 | 0.161 |
| | (0.19) | (0.17) | (0.23) | (0.23) |
| Female president | 0.014 | -0.009 | -0.036 | -0.054 |
| | (0.05) | (0.04) | (0.05) | (0.05) |
| Share of female candidates | | 0.642*** | 0.515*** | 0.620*** |
| | | (0.07) | (0.16) | (0.16) |
| H-index of female candidate is max in contest | | 0.218*** | 0.267*** | 0.286*** |
| | | (0.03) | (0.04) | (0.04) |
| Professorship | | | | 0.154*** |
| | | | | (0.05) |
| Gender × Discipline Fixed Effects | No | No | Yes | Yes |
| Gender × University Fixed Effects | No | No | Yes | Yes |
| Observations | 455 | 455 | 359 | 359 |

$*p < 0.10, **p < 0.05, ***p < 0.01$. The gender variable is coded as 1 for a woman and 0 for a man. Contests with only members of a specific gender are dropped in columns three and four.

# B Correlations between share of female jurors and probability that women are first-ranked

In this section I estimate the correlation between the share of women in committees and the ranks of women. These results ignore the possible endogeneity between the share of female jurors and unobserved components of candidate quality, and thus can only be correlational.

In the first table I look at the effect of two variables, the gender of the committee president and the proportion of females in the committee on the probability that the first-ranked candidate is a woman, using a probit regression. The results are in Table B.5. The first column has no controls, and there is a significant, positive effect of having more women in the committee in this case. However, once we control for the percentage of female candidates, this effect vanishes, and the coefficient becomes small and insignificant.

As in the main text of the article, I move to results at the candidate level. The results in the first three columns of Table B.6 can be interpreted as the effect of the feminisation of committees on the likelihood that a woman is first ranked. The correlations here are negative and significant, (i.e. more feminised committees are less likely to hire women) but controlling for field fixed effects and connections make the coefficient statistically insignificant.

Although we can likely rule out a large positive correlation of the feminisation of committees on the probability of women being hired, this doesn't preclude finding an effect on the *ranking* of women. I directly analyse ranks using the rank-ordered logit, in columns 4-6 of Table B.6. With this method I can only control for variables that affect ranking within committees. For all non-individual variables, I consider their effect on the rank of women relative to men. Similar to the results above, we cannot estimate these parameters in committees where only men or only women applied. Therefore, we have less observations than in the descriptive statistics presented in section 2.

A first interesting result is to look at the effect of being a woman on the rank given by the committee for both specifications, in columns 1 and 4. This effect is negative and significant even when controlling for the h-index. Though this article does not focus on discrimination *per se*, this column is the closest to a direct test since it asks whether women are as likely to be hired as men with the same publication record within a given committee. Next, consider the

Table B.6: Correlation between the gender of jurors and the gender of first ranked candidates

| Dependent variable: | Ranked first | | | Rank | | |
|---|---|---|---|---|---|---|
| Estimator: | C. Logit | C. Logit | C. Logit | RO. Logit | RO. Logit | RO. Logit |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| *Interacted with gender of the candidate* | | | | | | |
| **Share of women in the committee** | | -0.659* | -0.072 | | -0.403** | 0.252 |
| | | (0.35) | (1.32) | | (0.19) | (0.81) |
| Female president | | -0.381 | -0.542* | | -0.050 | -0.091 |
| | | (0.29) | (0.32) | | (0.16) | (0.18) |
| Share of female candidates | | | 0.307 | | | -0.318 |
| | | | (0.83) | | | (0.44) |
| Professorship | | | 0.894*** | | | 0.347* |
| | | | (0.32) | | | (0.19) |
| *Variable at the candidate level* | | | | | | |
| Candidate is a woman | -0.417*** | | | -0.212*** | | |
| | (0.14) | | | (0.08) | | |
| Standardised age-adjusted H-index | 0.326*** | 0.320*** | 0.355*** | 0.275*** | 0.275*** | 0.287*** |
| | (0.07) | (0.07) | (0.08) | (0.04) | (0.04) | (0.05) |
| PhD supervisor in the committee | | | 0.466* | | | 0.351** |
| | | | (0.26) | | | (0.15) |
| PhD from the same institute | | | 0.576*** | | | 0.295*** |
| | | | (0.17) | | | (0.10) |
| Gender × Discipline Fixed Effects | No | No | Yes | No | No | Yes |
| Gender × University Fixed Effects | No | No | Yes | No | No | Yes |
| Observations | 1357 | 1357 | 1357 | 1357 | 1357 | 1357 |

$*p < 0.10, **p < 0.05, ***p < 0.01$. The gender variable is coded as 1 for a woman and 0 for a man. Variables interacted with the gender of the candidate should be interpreted as the differential effect of the variable for women with respect to men. Committees where only one candidate was ranked, or where only applicants of a specific gender applied are dropped from the analysis.

interaction between how women are ranked, and the gender composition of the committee. There is a significant and large negative correlation between the share of women in the committees on the ranking of women in columns 2 and 5, where the only control is the h-index of candidates. However, once more control variables are added, this effect becomes insignificant, and is also likely driven by the effect of the reform. Results using the random utility model are very similar and are presented in Table B.7.

## B.1 Investigating endogeneity concerns

One important issue is that the selection of candidates into the candidate pool is possibly endogenous. For instance, candidates may put less effort in their application depending on the composition of the jury, or even decide not to apply at all. I investigate this concern by considering whether the composition of the jury affects variables other than the ranking within field. In the following tables, I present results both with contested and uncontested (only candidates of a single gender applied) committees. In Table B.8, I regress the gender ratio of candidates on the gender ratio of jury members, while controlling for discipline fixed-effects.

The applicant pool depends significantly on the gender composition of the committee in three of the estimations, either through sub-discipline effects (i.e. sub-disciplines within each field have different gender compositions) or through endogenous selection of candidates (candidates observe the gender composition of the committee and then decide whether to apply or not), though the effect becomes insignificant once I control for Gender × Discipline Fixed Effects and drop uncontested committees. We may therefore be worried that the results presented above could be endogenous, since candidate pools may have selection bias on unobservable candidate quality.

Table B.7: Correlation between the gender of jurors and the probability that a woman is better ranked within a dyad

| Dependent variable: | Probability that W is more highly ranked | | | |
|---|---|---|---|---|
| Estimator: | Probit | Probit | Probit | Probit |
| Marginal effects | (1) | (2) | (3) | (4) |
| Share of women in the committee | -0.092 | -0.036 | -0.065 | -0.071 |
| | (0.20) | (0.19) | (0.21) | (0.21) |
| Female President | -0.022 | -0.023 | -0.037 | -0.034 |
| | (0.05) | (0.04) | (0.05) | (0.05) |
| Standardised h-index difference | | 0.084*** | 0.087*** | 0.087*** |
| | | (0.01) | (0.01) | (0.01) |
| Share of female candidates | | | -0.100 | -0.115 |
| | | | (0.18) | (0.19) |
| Professorship | | | 0.083* | 0.088* |
| | | | (0.05) | (0.05) |
| PhD supervisor is in the committee | | | | 0.061 |
| | | | | (0.05) |
| PhD from the same institute | | | | 0.044 |
| | | | | (0.03) |
| Gender × Discipline Fixed Effects | No | No | Yes | Yes |
| Gender × University Fixed Effects | No | No | Yes | Yes |
| Observations | 1086 | 1086 | 1086 | 1086 |

$*p < 0.10, **p < 0.05, ***p < 0.01$. Dyads represent pairs of female-male candidates ranked by the same jury. The gender variable is coded as 1 for a woman and 0 for a man. Standard errors are clustered at the jury level. Variables on the PhD's of candidates are coded as follows: If both or neither candidates have an advisor in the committee, then the variable is equal to 0. If a candidate has an advisor in the committee and not the other, then the variable is coded either as 1 or -1, depending on the gender of the connected candidate.

Table B.8: Regression of the gender composition of applicant pools on the gender composition of the recruitment committee

| Dependent variable: | Gender composition of applicant pool | | | |
|---|---|---|---|---|
| Estimator | OLS | OLS | Within | Within |
| Marginal effects | (1) | (2) | (3) | (4) |
| Share of women in the committee | 0.674*** | 0.534*** | 0.175* | 0.061 |
| | (0.11) | (0.09) | (0.10) | (0.07) |
| Female president | 0.037 | 0.009 | 0.010 | -0.013 |
| | (0.03) | (0.02) | (0.02) | (0.02) |
| Gender × Discipline Fixed Effects | No | No | Yes | Yes |
| Observations | 455 | 359 | 455 | 359 |

$*p < 0.10, **p < 0.05, ***p < 0.01$. The gender variable is coded as 1 for a woman and 0 for a man. Standard errors are clustered at the jury level. Columns 2 and 4 drop contests with only male or female applicants.

# C Jury member fixed effects

Since decisions are made conjointly with other members of the jury, it is complicated to add jury fixed effects since the decision of the jury cannot be attributed to a single jury member. What we can do is look at the decisions of juries a particular professor was a member of, ascribe the decision of the jury to that member, and then see how the reform affects the decisions. Some members sit on multiple juries, and we can therefore add jury fixed effects to capture their mean gender bias, and see how it varies with the effect of the reform. This method has a few problems, one of which is that it changes the sample of relevant committees quite drastically; some committees will never appear since they don't have a jury member that appears in another committee, others will be repeated twice if they have many popular jurors. It is also complicated to understand jury fixed effects in this context, since individual jurors only account for 1/10th of the committee on average.[42] This method is only included to investigate juror fixed effects in this context and the results should be interpreted with caution.

With this retooled data set, each observation is equal to a one jury sitting in a particular committee, so that a committee with 16 members will appear 16 times. My outcome variable is the probability that a woman is ranked first by the committee. I first look at the endogenous gender bias in this data set, i.e. whether female jurors are less likely to rank female candidates first, and then analyse the effect of the reform, interacting with the gender of jury members.

The results are presented below in Table C.9. There is no raw gender bias, as can be seen in column 1, and the effect of the reform is resistant to jury fixed effects as can be seen in column 2. If we interact the gender of the jury members with the dummy variables for the reform and the treatment group as in columns 3 and 4, the results are similar to those presented by interacting the effect of the reform with the gender of the jury president; the change in discrimination seems to be driven by male jury members, though the difference between men and women are not significant. If we restrict the sample to jury members who have sat more than three times in the sample, as in columns 5 and 6, then these differences are significant, with the same sign as the jury president fixed effects.

# D Random utility model

The first models using a latent utility framework date back to Thurstone (1927). In those early models, ranked data was analysed by pairwise difference. Using these pairwise differences we can estimate the effect of the quota in terms of the increase in the share of female committee members.

Consider the model described in Section 3. We have a latent utility function where the candidate $i$ has a better ranking than candidate $i'$ if

$$U_{ij} \geq U_{i'j}$$

This implies

$$q_i \geq q_{i'} - \mu g_i g_j + \mu g_{i'} g_j$$

Assume that $q_i$, the candidate's intrinsic quality as evaluated by jury members depends on two factors: an observed term, $x_i$, and an unobserved term, $\epsilon_{ij} \sim N(0, \sigma)$ which represents other variables that the jury may take into account. If we represent the probability of candidate $i$ receiving a better ranking than candidate $i'$ by a binary variable $Y$, then:

$$Pr(Y = 1|x) = Pr(U(i) \geq U(i'))$$

$$Pr(x_i\beta + \mu g_i g_j + \epsilon_{ij} \geq x_{i'}\beta + \mu g_{i'} g_j + \epsilon_{i'j})$$

$$Pr(\beta(x_i - x_{i'}) + \mu(g_i - g_{i'})g_j \geq \epsilon_{i'j} - \epsilon_{ij})$$

[42] However, the jurors who appear many times may have an outsize influence in committee decisions.

Table C.9: Effect of the reform with jury fixed effects

| Dependent variable: | First ranked candidate is a woman | | | | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Estimator | OLS | OLS | OLS | OLS | OLS | OLS |
| Female president | 0.001 | -0.043 | -0.063 | -0.058 | -0.041 | -0.067 |
| | (0.05) | (0.03) | (0.06) | (0.04) | (0.07) | (0.06) |
| Female juror | 0.003 | 0.000 | 0.013 | 0.000 | -0.023 | 0.000 |
| | (0.01) | (.) | (0.03) | (.) | (0.06) | (.) |
| H-index of female candidate is max in contest | 0.309*** | 0.305*** | 0.295*** | 0.335*** | 0.230*** | 0.264*** |
| | (0.06) | (0.03) | (0.06) | (0.03) | (0.08) | (0.05) |
| Share of female candidates | 0.585*** | 0.644*** | 0.892*** | 0.899*** | 0.818*** | 0.847*** |
| | (0.13) | (0.11) | (0.19) | (0.10) | (0.24) | (0.15) |
| **Post Reform × Treatment Group** | | -0.209*** | -0.227** | -0.239*** | -0.295** | -0.334*** |
| | | (0.06) | (0.11) | (0.07) | (0.15) | (0.12) |
| Treatment Group | | 0.113 | 0.225 | 0.176 | 0.274 | -0.143 |
| | | (0.08) | (0.15) | (0.16) | (0.22) | (0.19) |
| Post-Reform | | 0.133*** | 0.154** | 0.118** | 0.160 | 0.226*** |
| | | (0.04) | (0.08) | (0.06) | (0.10) | (0.08) |
| **Post Reform × Treatment Group** | | | 0.040 | 0.067 | 0.275* | 0.430* |
| **× Female juror** | | | (0.06) | (0.12) | (0.15) | (0.22) |
| Treatment Group × Female juror | | | -0.045 | 0.021 | -0.070 | 0.098 |
| | | | (0.05) | (0.14) | (0.09) | (0.25) |
| Post-Reform × Female juror | | | -0.029 | 0.041 | -0.065 | -0.042 |
| | | | (0.04) | (0.08) | (0.08) | (0.11) |
| Professorship | | | 0.201*** | 0.198*** | 0.192** | 0.222*** |
| | | | (0.06) | (0.04) | (0.08) | (0.06) |
| Establishment Fixed Effects | No | No | Yes | Yes | Yes | Yes |
| Discipline Fixed Effects | No | No | Yes | Yes | Yes | Yes |
| Jury Fixed Effects | No | Yes | No | Yes | No | Yes |
| Observations | 4002 | 4002 | 1802 | 1802 | 534 | 534 |

$*p < 0.10, **p < 0.05, ***p < 0.01$. The gender variable is coded as 1 for a woman and 0 for a man. The treatment variable is based on mean pre-reform gender parity at the field*university level. Columns 5 and 6 only keep jurors who have sat more than 3 times in my sample.

$$= \Phi(X\beta + \mu G g_j)$$

where $X = x_i - x_{i'}$ and $G = g_i - g_{i'}$. We can then use a probit regression to recover the parameters. Of course, when $g_i = g_{i'}$, $G = 0$, so we can only estimate $\mu$ on pairs of candidates of different gender. Only dyads with candidates of different gender can give us information on own-gender preferences.[43]

With the theoretical justification set out above, we can estimate the dyadic model through IV, instrumenting for the potentially endogenous variable with the effect of being in our group post reform. I estimate the following equations:

$$W_j = \alpha^1 R_j T_j + \gamma^1 T_j + \delta^1 R_j + \beta^1 (X_i - X_{i'}) + \beta_2^1 Y_j + \epsilon_{ij}^1$$

$$\mathrm{D}_{ii'j} = \mu \hat{W}_j + \gamma^2 T_j + \delta^2 R_j + \beta_1^2 (X_i - X_{i'}) + \beta_2^2 Y_j + \epsilon_{ij}^2$$

where $D_{ii'j}$ is a binary variable equal to 1 if the female candidate i has a better rank than the male candidate i' within a dyad, $(X_i - X_{i'})$ is a vector of differences between candidate level characteristics, for instance the difference in h-index between candidates i and i', $\hat{W}$ is the instrumented share of women in committees, and $\mu$ is our coefficient of interest, estimating the effect of having more women in recruitment committees on the probability of being hired. The model is estimated through 2SLS.

The first-stage estimates are presented in Table D.10.[44] Being in the group affected by the

[43] Simulations, presented in Section G show that these estimates are unbiased for the data generating process presented above, and that the standard errors from OLS (if the parameter values are not too large) or probit regressions have the right power once we cluster at the committee level, to take correlation of errors (since we may use the same individual from the same committee multiple times in our sample) into account.

[44] The F-test statistic for the first-stage is above 10 (and in fact close to 30) in all columns, as advised by Staiger and Stock (1997), and the instrument is significantly correlated with the endogenous variable.

Table D.10: First stage: IV

| Dependent variable: | Share of women in committee | | |
|---|---|---|---|
| Estimator: | IV | IV | IV |
| | (1) | (2) | (3) |
| **Post Reform × Treatment Group** | 0.174*** | 0.173*** | 0.170*** |
| | (0.03) | (0.03) | (0.03) |
| Female President | 0.027** | 0.027** | 0.023* |
| | (0.01) | (0.01) | (0.01) |
| Treatment Group | -0.197*** | -0.196*** | -0.153*** |
| | (0.03) | (0.03) | (0.04) |
| Post-Reform | -0.048*** | -0.048*** | -0.046** |
| | (0.02) | (0.02) | (0.02) |
| Standardised h-index difference | | -0.004 | -0.003 |
| | | (0.00) | (0.00) |
| Professorship | | | -0.033** |
| | | | (0.01) |
| *Differenced variables* | | | |
| PhD supervisor is in the committee | | -0.004 | -0.002 |
| | | (0.01) | (0.01) |
| PhD from the same institute | | 0.003 | 0.001 |
| | | (0.01) | (0.01) |
| Share of female candidates | | | 0.030 |
| | | | (0.04) |
| F-Statistic | 35.83 | 36.04 | 33.02 |
| Observations | 1086 | 1086 | 1086 |

$*p < 0.10, **p < 0.05, ***p < 0.01$. Dyads represent pairs of female-male candidates ranked by the same jury. Robust standard errors in parentheses. The gender variable is coded as 1 for a woman and 0 for a man. Variables on the PhD of candidates are coded as follows: If both or neither candidates have an advisor in the committee, then the variable is equal to 0. If a candidate has an advisor in the committee and not the other, then the variable is coded either as 1 or -1, depending on the gender of the connected candidate.

reform has a highly significant and positive (around 17 percentage points) effect on the share of women in committees, as expected. There is also a negative effect of being in the control group on this share, which is significant at the 5 percent level. The instrumental variable results are presented in Table D.11. The share of women in the committee has a significant negative effect. The effect size can now be interpreted in terms of changes in the share of women in committees. The point estimates for column 2 imply that a 10% increase in the percentage of women in committees has the same effect on the probability that a woman is more highly ranked in a dyad than a 3 standard deviation decrease in the h-index.

# E Full effect of the reform

In this section I look at the direct effect of the reform without considering heterogeneity in treatment effects across disciplines, using only a dummy variable for the effect of the reform. With this methodology, the net effect of the reform seems to be close to 0. The estimate is insignificant, and the point estimate of the effect is 4 times smaller than that of a standard deviation in age-discounted h-index.

# F Control Function Approach

Another way to deal with endogeneity in estimation is to use a control function approach. This entails regressing the endogenous variable on the instrument in the first stage, then including the residuals and the endogenous variable in the main specification. One benefit is that this approach can be used to accommodate non-linear models. The results can be seen in Table F.13. The main results are only significant at the ten percent level in this specification. The point estimates can be interpreted as in the main part of the text, by comparing with the size of the

Table D.11: IV estimate of the effect of the increase in women jurors on the probability that a woman is better ranked within a dyad

| Dependent variable: | Woman is better ranked within a dyad | | |
|---|---|---|---|
| Estimator: | IV | IV | IV |
| | (1) | (2) | (3) |
| **Share of women in the committee** | -2.017*** | -2.017*** | -1.813** |
| | (0.64) | (0.63) | (0.71) |
| Treatment Group | -0.139** | -0.136** | 0.161 |
| | (0.07) | (0.07) | (0.13) |
| Post-Reform | 0.013 | 0.011 | 0.019 |
| | (0.06) | (0.06) | (0.06) |
| Female President | 0.030 | 0.028 | -0.000 |
| | (0.05) | (0.05) | (0.05) |
| Share of female candidates | | | -0.023 |
| | | | (0.21) |
| *Differenced variables* | | | |
| Standardised h-index difference | | 0.072*** | 0.075*** |
| | | (0.01) | (0.01) |
| PhD supervisor is in the committee | | 0.056 | 0.066 |
| | | (0.05) | (0.05) |
| PhD from the same institute | | 0.048 | 0.050 |
| | | (0.04) | (0.03) |
| Discipline Fixed Effects | No | No | Yes |
| University Fixed Effects | No | No | Yes |
| Observations | 1086 | 1086 | 1086 |

$*p < 0.10, **p < 0.05, ***p < 0.01$. Dyads represent pairs of female-male candidates ranked by the same jury. The treatment variable is based on mean pre-reform gender parity at the field*university level. The gender variable is coded as 1 for a woman and 0 for a man. Errors are clustered at the contest level.

Table E.12: Global effect of the reform

| Dependent variable: | Rank | | |
|---|---|---|---|
| Estimator: | RO. Logit | RO. Logit | RO. Logit |
| | (1) | (2) | (3) |
| *Interacted with gender of the candidate* | | | |
| Post-Reform | -0.048 | -0.063 | -0.101 |
| | (0.15) | (0.15) | (0.19) |
| Share of female candidates | | -0.302 | -0.367 |
| | | (0.39) | (0.44) |
| *Variable at the candidate level* | | | |
| Candidate is a woman | -0.213** | -0.031 | |
| | (0.11) | (0.22) | |
| Standardised age-adjusted H-index | | 0.275*** | 0.284*** |
| | | (0.04) | (0.05) |
| PhD supervisor in the committee | | | 0.356** |
| | | | (0.15) |
| PhD from the same institute | | | 0.287*** |
| | | | (0.10) |
| Gender × Discipline Fixed Effects | No | No | Yes |
| Gender × University Fixed Effects | No | No | Yes |
| Observations | 1357 | 1357 | 1357 |

$*p < 0.10, **p < 0.05, ***p < 0.01$. The gender variable is coded as 1 for a woman and 0 for a man. Variables interacted with the gender of the candidate should be interpreted as the differential effect of the variable for women with respect to men.

Table F.13: Control Function Approach

| Dependent variable: | Rank | | |
|---|---|---|---|
| Estimator: | RO. Logit | RO. Logit | RO. Logit |
| | (1) | (2) | (3) |
| *Interacted with gender of the candidate* | | | |
| Share of women in committee | -5.293* | -6.003* | -4.719 |
| | (3.07) | (3.23) | (3.57) |
| Female president | 0.050 | 0.033 | -0.030 |
| | (0.25) | (0.26) | (0.30) |
| Treatment Group | -0.367 | -0.326 | 0.750 |
| | (0.34) | (0.32) | (2.35) |
| Post-Reform | 0.031 | -0.017 | 0.002 |
| | (0.25) | (0.27) | (0.31) |
| Share of female candidates | | 0.329 | -0.093 |
| | | (0.67) | (0.70) |
| Professorship | | 0.268 | 0.195 |
| | | (0.29) | (0.33) |
| *Variable at the candidate level* | | | |
| Residuals | 6.013* | 7.360** | 5.952 |
| | (3.29) | (3.42) | (3.84) |
| Standardised age-adjusted H-index | | 0.300*** | 0.294*** |
| | | (0.06) | (0.06) |
| PhD supervisor in the committee | | 0.439* | 0.425* |
| | | (0.24) | (0.24) |
| PhD from the same institute | | 0.292** | 0.309** |
| | | (0.14) | (0.15) |
| Candidate is a woman | 2.360 | 2.441* | 2.198 |
| | (1.53) | (1.48) | (1.69) |
| Gender × Discipline Fixed Effects | No | No | Yes |
| Gender × University Fixed Effects | Yes | Yes | Yes |
| Observations | 1357 | 1357 | 1357 |

$*p < 0.10, **p < 0.05, ***p < 0.01$. Standard errors estimated via bootstrap. The gender variable is coded as 1 for a woman and 0 for a man. Variables interacted with the gender of the candidate should be interpreted as the differential effect of the variable for women with respect to men.

coefficient for the standardised h-index. An increase in the number of women in committees by 10 percent has approximately the same effect on the rank of women by the same amount as a 2 standard deviation drop in the h-index.

# G Differences between the Random Utility Model and the Rank-Ordered Logit

Are the two methods presented in the model section of the article equivalent? Consider the following data generating process. I generate a jury utility variable $U_{ij}$ such that:

$$U_{ij} = \beta q_{ij} + \mu g_i g_j + \epsilon_{ij}$$

where $g_i$ and $g_j$ are binomial variables, $q_{ij}$ is normally distributed with mean 5 and standard deviation of 1. Using the values from the article, I use 0.05 for $\beta$ and -0.6 for $\mu$.

From this utility variable I then simulate artificial committees in the following way: from a 1000 observations of $U_{ij}$ I create 100 committees with 10 candidates in each committee. 60 committees have a female jury president. Each committee then creates a ranking $r_{ij}$ of each

Table G.14: Power of estimation methods

| Estimator | RO Logit | OLS | Probit | Logit |
|---|---|---|---|---|
| *Normal errors* | | | | |
| Mean | -0.03 | -0.00 | -0.00 | 0.00 |
| SD | 0.11 | 0.04 | 0.04 | 0.04 |
| # of t-statistics $\leq$\|1.96\| (%) | 947 (95%) | 941(94%) | 941(94%) | 941(94%) |
| *Type I Extreme Value* | | | | |
| Mean | 0.00 | 0.00 | 0.00 | 0.00 |
| SD | 0.11 | 0.04 | 0.04 | 0.04 |
| # of t-statistics $\geq$\|1.96\| (%) | 943 (94%) | 960(96%) | 960(96%) | 959(96%) |

Results from 1000 simulations. True effect is 0. For the right amount of power, tests should reject the null hypothesis only 5% of the time. Results from the probit and logit estimations are marginal effects.

Table G.15: Estimates from simulations

| Estimator | RO Logit | OLS | Probit | Logit |
|---|---|---|---|---|
| *Gender bias* | | | | |
| Mean | -0.61 | -0.15 | -0.38 | -0.60 |
| SD | 0.11 | 0.04 | 0.10 | 0.17 |
| *Effect of quality* | | | | |
| Mean | 0.05 | 0.01 | 0.03 | 0.05 |
| SD | 0.04 | 0.01 | 0.03 | 0.04 |
| *Ratio* | | | | |
| Median | -0.086 | -0.086 | -0.85 | -0.085 |
| SD | 0.07 | 0.10 | 0.10 | 0.10 |

Means and standard deviation of estimators of $\mu$ and $\beta$ from 1000 simulations. Error term is distributed type I EV. Results are similar with normally distributed errors. Results from the probit and logit estimations are full effects and not marginal. The marginal effects are similar in size to the OLS results.

candidate based on the value of $U_{ij}$.

This simulation allows me to answer two different questions: do both methods have the right power, and can both methods accurately recover the $\mu/\beta$ ratio?

Let us consider the first question. The right standard errors would give us a 5% probability of having a t-statistic above the absolute value of 1.96, if the true effect was in fact 0. I simulate the DGP described above 500 times, with a true effect of gender bias, $\mu$ equal to 0. I present results both with $\epsilon_{ij}$ either normally distributed or type I extreme value in Table G.14. In all cases, errors are clustered at the contest level, since candidates evaluated by the same jury are compared multiple times, generating correlation between errors.

The methods that use the random utility approach are very similar in terms of power and estimates. They seem to over-reject when errors are distributed Extreme Value, and under-reject when errors are distributed normally. However, the power estimates are within 1% of the valid rejection rates. The rank-ordered logit has a wider standard deviation of estimates, but is closer to the valid rejection rates.

In Table G.15, we can see the mean values of the estimators for the parameters presented above. I present the median value for the $\mu/\beta$ ratio, since the mean is affected by outliers, i.e. when $\beta$ is estimated to be very close to 0. The mean point estimates are similar for the rank-ordered logit, probit and OLS. The results are similar no matter the method considered.

This is the case when the values of the parameters are small. However, with larger values of $\beta$ and $\mu$, the LPM suffers from bias, as can be seen in Table G.16. This well known result stems from the predicted values of Y, $X\beta + \mu G_{ij}$, being outside the range of possible values that can be taken on by a binary variable, i.e. 0-1.

Table G.16: Estimates from simulations: large parameters

| Estimator | RO. Logit | OLS | Probit | Logit |
|---|---|---|---|---|
| *Gender bias* | | | | |
| Mean | -6.01 | -0.40 | -3.33 | -6.04 |
| SD | 0.30 | 0.02 | 0.20 | 0.36 |
| *Effect of quality* | | | | |
| Mean | 3.00 | 0.22 | 1.66 | 3.01 |
| SD | 0.13 | 0.01 | 0.08 | 0.14 |
| *Ratio* | | | | |
| Median | -0.50 | -0.55 | -0.50 | -0.50 |
| SD | 0.02 | 0.02 | 0.02 | 0.02 |

Means and standard deviation of estimators of $\mu$ and $\beta$ from 1000 simulations. Error term is distributed type I EV. Results are similar with normally distributed errors. Results from the probit and logit estimations are full effects and not marginal. The marginal effects are similar in size to the OLS results.

# H Additional tables

Figure H.1 plots the difference between control and treatment group, with the standard errors on the difference.

To see whether publications could explain the gender difference in hiring, I look at gender differences in h-indices and Citations in Table H.17, regressing 4 measures of publication on a set of fixed effects, which include whether the post is for a full professorship, field fixed effects, and interactions between field and professorship. To account for the large field differences in citations, I standardise the h-index measures by field in the main parts of the estimation. This variable is included in column 4. Female candidates have lower measures of publication quality no matter the measure chosen, though the effect on citations is not significant. It is therefore crucial to control for differences in publications in the estimation.

I analyse whether the reform affected gender differences in publications in Table H.18. I control for contest fixed effects to compare candidates who apply for the same position, using two different measures of publications. There is no effect of the reform on differences in publications, and no differential effect of the reform in the treatment group relative to the control group

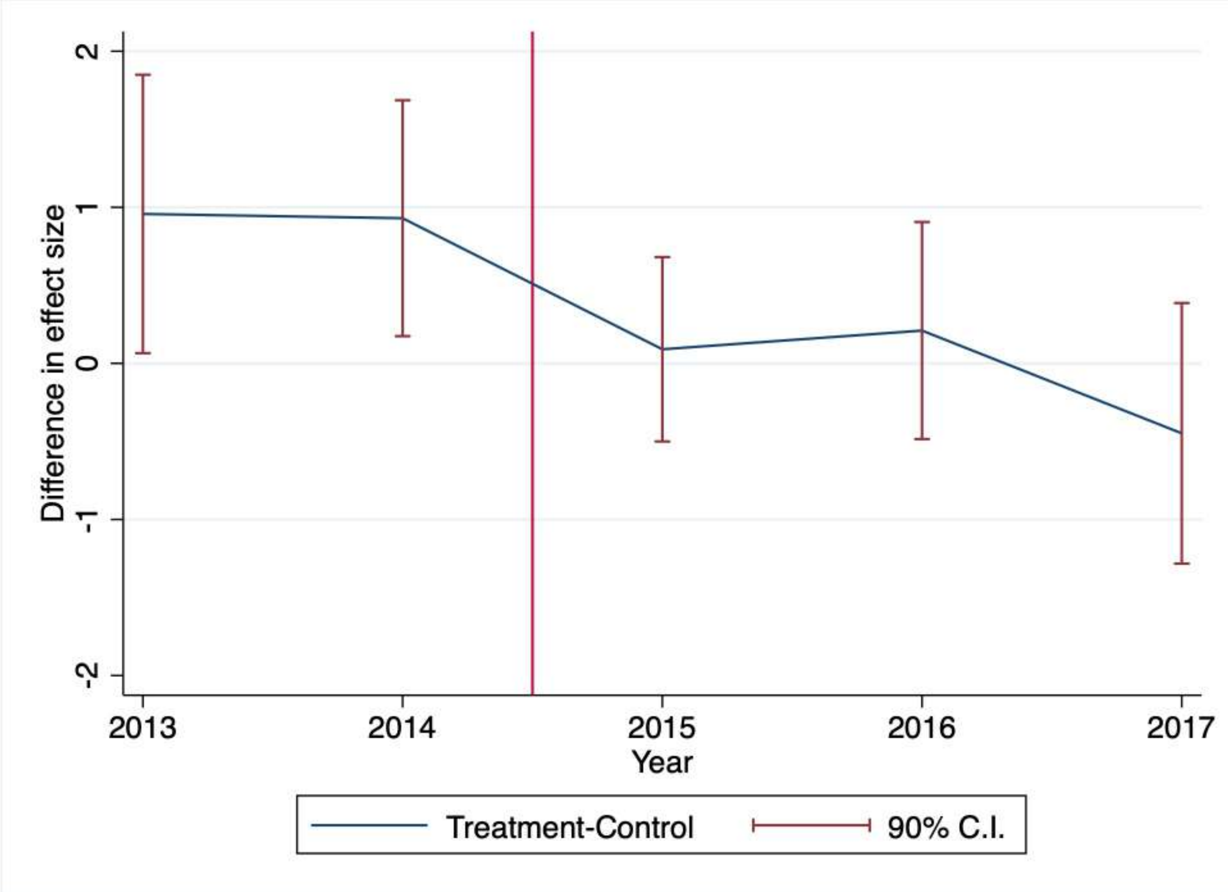
In Table H.19, I rerun the analysis presented in Table 9 at the contest level. The results are similar: male presidents seem to become much harsher after the reform in the treatment group, with no significant change for female presidents in the treatment group or for male presidents in the control group.

In Table H.20, I use the same specification as in Table 2, but include other measures of publications that could be more relevant than the h-index. I include variables with age- and author-discounted h-indices and citations. The variable that seems most significant in this case the age-discounted h-index, which is why I choose this variable in the main specifications of the article as a control for publications. This table include the h-index of the jurors, and the mean age of the jury members, for which I have data for a subset of the data only. Including these other citation measures makes no difference to the effect of the reform. There is no effect of the mean jury h-index, however older committees seem to view women more positively.

In Table H.21, I look at how the share of women among ranked candidates evolves with the reform. I use the same methodology as in Table 5. Here there seems to be a slight concern since women are more likely to be ranked after the reform in the treatment group, though the difference between treatment and control group is not significant.

In Table H.22, I try to see whether dropping lower ranked candidates can help mitigate these concerns. This is the sample from columns 3 and 4 Table 6. With this specification, we see that the point estimate for the change between treatment and control group following the reform is

Figure H.1: Difference in the effect of gender on ranks between treatment and control group




much smaller than in the regular sample, especially once we consider only the two years either side of the reform. Since this sample is actually the one that gives the highest point estimates in the article, it seems unlikely that changes in selection into the ranking pool drives the results.

We can look at whether jury characteristics, such as the average h-index of jury members and the age and h-indices of the jury presidents,[45] affect the probability of a woman being hired. I estimate 4 models in Table H.23 that include these variables into the rank-ordered and conditional logit. In the standard model that does not look at the effect of the reform, none of these variables have a significant effect. Two variables are significant at the 10% level in the rank-ordered logit however once I include these variables in the model that looks at the effect of the quota. Older presidents, and committees where the male committee members have high h-indices seem to favour women.[46]

In Table H.24, I test for pre-trends between the treatment and the control group, with the three methods used in the main part of the article (conditional logit, rank-ordered logit, probit on dyads). I find no significant trend, though the point estimates are high.

In Table H.25, I keep only labs where the mean share of women in committees before the reform belongs to the interval ]0.33,0.47[, excluding all labs that were far away from the threshold and which may thus be too dissimilar to each other. This reduces my sample size to 553 observations only. The results without field fixed effects are still significant at the 5% level, but become insignificant once field fixed effects are added. The point estimates remain high however and are if anything slightly higher than in the main part of the text.

In Table H.26, I add year $\times$ gender fixed effects to the main specification presented in 2. The

[45]For a few presidents I was not able to recover the age of the presidents. This leads to the change in the number of observations.

[46]The interactions of these variables with the effect of the reform are not significant at all however.

Table H.17: Gender differences in h-indices

| Dependent variable: | H-index | Age-adjusted h-index | Citations | Standardised age-adjusted h-index |
|---|---|---|---|---|
| Estimator: | OLS | OLS | OLS | OLS |
| Candidate is a woman | -1.152*** | -0.186 | -49.856 | -0.131** |
| | (0.26) | (0.12) | (50.02) | (0.06) |
| Fixed effects | Yes | Yes | Yes | Yes |
| Observations | 1548 | 1548 | 1548 | 1548 |

$*p < 0.10, **p < 0.05, ***p < 0.01$. The gender variable is coded as 1 for a woman and 0 for a man. Fixed effects include University, whether the job is for a professorship, field and field × professorship interaction terms.

Table H.18: Gender differences in h-indices

| Dependent variable: | Age-adjusted standardised h-index | | Unadjusted h-index | |
|---|---|---|---|---|
| Estimator: | OLS | OLS | OLS | OLS |
| *Interacted with gender of the candidate* | | | | |
| Post-Reform | 0.026 | 0.031 | 0.023 | 0.261 |
| | (0.12) | (0.15) | (0.56) | (0.73) |
| Treatment Group | | 0.050 | | 0.388 |
| | | (0.18) | | (0.87) |
| **Post Reform × Treatment Group** | | -0.015 | | -0.577 |
| | | (0.23) | | (1.13) |
| *Variable at the candidate level* | | | | |
| Candidate is a woman | -0.113 | -0.134 | -0.908** | -1.066* |
| | (0.09) | (0.12) | (0.43) | (0.56) |
| Contest level fixed effects | Yes | Yes | Yes | Yes |
| Observations | 1357 | 1357 | 1357 | 1357 |

$*p < 0.10, **p < 0.05, ***p < 0.01$. The gender variable is coded as 1 for a woman and 0 for a man. Fixed effects are at the contest level

results are very close to those presented in the main part of the text.

Table H.19: Probit on gender of first-ranked candidate

| Dependent variable: | Gender of first-ranked candidate | |
|---|---|---|
| Estimator: | Probit | Probit |
| Marginal Effects | (1) | (2) |
| **Post Reform × Treatment Group** | -0.263*** | -0.333** |
| | (0.09) | (0.13) |
| **Post Reform × Treatment Group** | 0.349 | 0.492* |
| **× Female President** | (0.22) | (0.30) |
| Treatment Group | 0.170*** | 0.296*** |
| | (0.06) | (0.11) |
| Treatment Group | -0.228* | -0.347* |
| × Female President | (0.14) | (0.20) |
| Post-Reform | 0.130** | 0.196** |
| | (0.06) | (0.09) |
| Post-Reform | -0.125 | -0.185 |
| × Female President | (0.12) | (0.17) |
| H-index of female candidate is max in contest | 0.229*** | 0.284*** |
| | (0.03) | (0.03) |
| Share of female candidates | 0.642*** | 0.502*** |
| | (0.04) | (0.14) |
| Female President | 0.071 | 0.084 |
| | (0.07) | (0.10) |
| Gender × University Fixed Effects | Yes | Yes |
| Gender × Discipline Fixed Effects | No | Yes |
| Observations | 455 | 359 |

$*p < 0.10, **p < 0.05, ***p < 0.01$. The gender variable is coded as 1 for a woman and 0 for a man. This regression considers first-ranked candidates only. Contests with applicants from a specific gender only are dropped in column four.

Table H.20: Rank-Ordered Logit using the quota: Other publication measures

| Dependent variable: | Rank | | |
|---|---|---|---|
| Estimator: | RO. Logit | RO. Logit | RO. Logit |
| | (1) | (2) | (3) |
| *Interacted with gender of the candidate* | | | |
| **Post Reform × Treatment Group** | -0.698** | -1.007*** | -0.815** |
| | (0.31) | (0.37) | (0.40) |
| Post-Reform | 0.228 | 0.253 | 0.175 |
| | (0.20) | (0.20) | (0.23) |
| Treatment Group | 0.360* | 0.693** | 1.493*** |
| | (0.22) | (0.32) | (0.58) |
| Female president | -0.077 | -0.051 | -0.108 |
| | (0.17) | (0.16) | (0.18) |
| Share of female candidates | | -0.166 | -0.206 |
| | | (0.41) | (0.43) |
| Professorship | | | 0.364* |
| | | | (0.19) |
| *Variable at the candidate level* | | | |
| Candidate is a woman | -0.291** | | |
| | (0.15) | | |
| H-index of the candidate | -0.016 | -0.017 | -0.014 |
| | (0.02) | (0.02) | (0.03) |
| Age-discounted h-index | 0.150*** | 0.150*** | 0.144*** |
| | (0.05) | (0.05) | (0.05) |
| Co-author discounted | 0.040 | 0.042 | 0.045 |
| h-index of the candidate | (0.04) | (0.04) | (0.04) |
| Citations of the candidate | -0.000 | -0.000 | 0.000 |
| | (0.00) | (0.00) | (0.00) |
| Co-author discounted | 0.000 | 0.000 | -0.000 |
| citations of the candidate | (0.00) | (0.00) | (0.00) |
| Age-Discounted citations | -0.001 | -0.001 | -0.001 |
| of the candidate | (0.00) | (0.00) | (0.00) |
| PhD supervisor in the committee | | | 0.389** |
| | | | (0.16) |
| PhD from the same institute | | | 0.293*** |
| | | | (0.10) |
| Gender × Discipline Fixed Effects | No | No | Yes |
| Observations | 1355 | 1355 | 1355 |

$*p < 0.10, **p < 0.05, ***p < 0.01$. The gender variable is coded as 1 for a woman and 0 for a man. The h-index variable and other citation measures can be interpreted as the effect of increasing the h-index for all candidates. Variables interacted with the gender of the candidate should be interpreted as the differential effect of the variable for women with respect to men.

Table H.21: Effect of the reform on the share of women among ranked candidates

| Dependent variable: | Gender composition of ranked candidates | | | | | |
|---|---|---|---|---|---|---|
| Years | 2009-2018 | | | 2013-2016 | | |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Post-Reform | 0.013 | 0.103** | 0.009 | 0.003 | 0.052 | 0.004 |
| | (0.04) | (0.05) | (0.04) | (0.05) | (0.05) | (0.05) |
| Professorship | -0.177*** | -0.100** | -0.143*** | -0.170*** | -0.054 | -0.125*** |
| | (0.05) | (0.04) | (0.03) | (0.06) | (0.05) | (0.04) |
| Treatment Group | | | -0.071 | | | -0.038 |
| | | | (0.06) | | | (0.07) |
| Post-Reform × Treatment Group | | | 0.078 | | | 0.027 |
| | | | (0.06) | | | (0.07) |
| Group | Control Group | Treatment Group | Both | Control Group | Treatment Group | Both |
| Discipline× Establishment Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 235 | 220 | 455 | 166 | 130 | 296 |

$*p < 0.10, **p < 0.05, ***p < 0.01$. The gender variable is coded as 1 for a woman and 0 for a man. The treatment variable is based on mean pre-reform gender parity at the field*university level. The last 3 columns are restrained to 2013-2016.

Table H.22: Effect of the reform on the share of women among ranked candidates- Top 3 Ranks only

| Dependent variable: | Gender composition of candidates | | | | | |
|---|---|---|---|---|---|---|
| Years | 2009-2018 | | | 2013-2016 | | |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Post-Reform | 0.010 | 0.041 | 0.012 | 0.009 | 0.036 | 0.021 |
| | (0.04) | (0.05) | (0.04) | (0.05) | (0.05) | (0.05) |
| Professorship | -0.116** | -0.097** | -0.115*** | -0.120** | -0.031 | -0.087** |
| | (0.05) | (0.05) | (0.03) | (0.06) | (0.06) | (0.04) |
| Treatment Group | | | -0.111* | | | -0.046 |
| | | | (0.06) | | | (0.08) |
| Post-Reform × Treatment Group | | | 0.022 | | | 0.001 |
| | | | (0.06) | | | (0.07) |
| Group | Control Group | Treatment Group | Both | Control Group | Treatment Group | Both |
| Gender × Discipline Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 180 | 172 | 352 | 120 | 108 | 228 |

$*p < 0.10, **p < 0.05, ***p < 0.01$. The gender variable is coded as 1 for a woman and 0 for a man. Only the top 3 candidates are considered. The treatment variable is based on mean pre-reform gender parity at the field*university level. The last 3 columns are restrained to 2013-2016.

Table H.23: Effect of jury characteristics

| Dependent variable: | Ranked first | | Rank | |
|---|---|---|---|---|
| Estimator | C. Logit | C. Logit | RO. Logit | RO. Logit |
| | (1) | (2) | (3) | (4) |
| *Interacted with gender of the candidate* | | | | |
| H-index of the president | 0.068 | -0.008 | 0.007 | -0.027 |
| | (0.14) | (0.15) | (0.09) | (0.10) |
| Age of the Jury President | 0.004 | 0.004 | 0.003 | 0.019* |
| | (0.01) | (0.02) | (0.01) | (0.01) |
| Mean h-index of the | 0.158 | 0.112 | -0.076 | -0.153 |
| women in the committee | (0.25) | (0.26) | (0.16) | (0.17) |
| Mean h-index of the | -0.086 | -0.028 | 0.212 | 0.319* |
| men in the committee | (0.27) | (0.28) | (0.17) | (0.19) |
| Share of female candidates | -0.559 | -0.435 | -0.505 | -0.440 |
| | (0.73) | (0.78) | (0.40) | (0.40) |
| **Post Reform × Treatment Group** | | -1.512*** | | -0.737** |
| | | (0.58) | | (0.31) |
| Post-Reform | | 0.795** | | 0.310 |
| | | (0.39) | | (0.21) |
| Treatment Group | | 0.842* | | 0.383* |
| | | (0.45) | | (0.23) |
| Female president | | -0.421 | | -0.043 |
| | | (0.33) | | (0.18) |
| *Variable at the candidate level* | | | | |
| Standardised age-adjusted h-index | 0.310*** | 0.317*** | 0.267*** | 0.269*** |
| | (0.07) | (0.07) | (0.04) | (0.04) |
| Candidate is a woman | | -0.616 | | -1.155* |
| | | (1.12) | | (0.62) |
| Observations | 1332 | 1332 | 1332 | 1332 |

$*p < 0.10, **p < 0.05, ***p < 0.01$. The gender variable is coded as 1 for a woman and 0 for a man. The treatment variable is based on mean pre-reform gender parity at the field*university level. Variables interacted with the gender of the candidate should be interpreted as the differential effect of the variable for women with respect to men. Contests where only one candidate was ranked, or where only applicants of a specific gender applied are dropped from the analysis.

Table H.24: Testing for pre-trends

| | Rank | First Ranked | Winner of Dyad |
|---|---|---|---|
| Estimator | R.O. Logit | C. Logit | Probit |
| *Interacted with gender of the candidate* | | | |
| Treatment Group | 0.924 | 0.876 | 0.279** |
| | (0.71) | (1.10) | ( 0.14) |
| Differential trend in Treatment Group | 0.511 | 0.620 | 0.066 |
| | (0.33) | (0.65) | (0.08) |
| Trend | -0.163 | 0.118 | 0.009 |
| | (0.27) | (0.51) | (0.06) |
| Female president | -0.273 | -0.527 | -0.075 |
| | (0.29) | (0.54) | (0.07) |
| Share of female candidates | 0.509 | 1.434 | 0.273 |
| | (0.66) | (1.55) | (0.26) |
| Professorship | 0.607** | 1.218** | 0.209*** |
| | (0.29) | (0.53) | (0.07) |
| *Variables at the candidate level* | | | |
| Standardised age-adjusted H-index | 0.269*** | 0.354*** | 0.063*** |
| | (0.06) | (0.12) | (0.02) |
| PhD supervisor in the committee | 0.637** | 0.699* | 0.184*** |
| | (0.28) | (0.42) | (0.08) |
| PhD from the same institute | 0.106 | 0.478 | 0.026 |
| | (0.19) | (0.30) | (0.05) |
| Gender × Discipline Fixed Effect | Yes | Yes | Yes |
| Gender × University Fixed Effect | Yes | Yes | Yes |
| Time Period | 2009-2014 | 2009-2014 | 2009-2014 |
| Observations | 555 | 555 | 407 |

$*p < 0.10, **p < 0.05, ***p < 0.01$. For columns 1 and 2, the gender variable is coded as 1 for a woman and 0 for a man. The h-index variable and other citation measures can be interpreted as the effect of increasing the h-index for all candidates. Variables interacted with the gender of the candidate should be interpreted as the differential effect of the variable for women with respect to men. For column 3, dyads represent pairs of female-male candidates ranked by the same jury. The gender variable is coded as 1 for a woman and 0 for a man. Standard errors are clustered at the jury level. Variables on the PhD's of candidates are coded as follows: If both or neither candidates have an advisor in the committee, then the variable is equal to 0. If a candidate has an advisor in the committee and not the other, then the variable is coded either as 1 or -1, depending on the gender of the connected candidate.

Table H.25: Effect of the quota on the rank of female candidates, restricted sample to ]0.33,0.47[ mean share of women in committees before the reform

| Dependent variable: | Rank | | |
|---|---|---|---|
| Estimator: | RO. Logit | RO. Logit | RO. Logit |
| | (1) | (2) | (3) |
| *Interacted with gender of the candidate* | | | |
| **Post Reform × Treatment Group** | -1.504** | -1.672** | -0.710 |
| | (0.74) | (0.76) | (0.86) |
| Post-Reform | -0.049 | -0.000 | -0.241 |
| | (0.29) | (0.29) | (0.36) |
| Treatment Group | 1.014* | 1.214** | 1.341** |
| | (0.60) | (0.61) | (0.65) |
| Female president | -0.185 | -0.221 | -0.443 |
| | (0.27) | (0.28) | (0.32) |
| Professorship | | | 0.515 |
| | | | (0.35) |
| Share of female candidates | | 0.263 | -0.085 |
| | | (0.89) | (0.96) |
| *Variables at the candidate level* | | | |
| Standardised age-adjusted H-index | | 0.255*** | 0.250*** |
| | | (0.07) | (0.07) |
| PhD supervisor in the committee | | | 0.818*** |
| | | | (0.28) |
| PhD from the same institute | | | 0.439** |
| | | | (0.18) |
| Gender × Discipline Fixed Effects | No | No | Yes |
| Gender × University Fixed Effects | Yes | Yes | Yes |
| Observations | 553 | 553 | 553 |

$*p < 0.10, **p < 0.05, ***p < 0.01$. The gender variable is coded as 1 for a woman and 0 for a man. Robust standard errors in parentheses. The treatment variable is based on mean pre-reform gender parity at the field*university level. Variables interacted with the gender of the candidate should be interpreted as the differential effect of the variable for women with respect to men. Contests where only one candidate was ranked, or where only applicants of a specific gender applied are dropped from the analysis.

Table H.26: Effect of the quota on the rank of female candidates, with gender×year fixed effects

| Dependent variable: | Rank | | |
|---|---|---|---|
| Estimator: | RO. Logit | RO. Logit | RO. Logit |
| | (1) | (2) | (3) |
| *Interacted with gender of the candidate* | | | |
| **Post Reform × Treatment Group** | -0.799** | -0.954** | -0.833** |
| | (0.34) | (0.38) | (0.41) |
| Post-Reform | 0.099 | 0.485 | 0.337 |
| | (0.28) | (0.30) | (0.33) |
| Treatment Group | 0.497* | 0.689** | 1.601*** |
| | (0.26) | (0.33) | (0.60) |
| Female president | -0.082 | -0.100 | -0.153 |
| | (0.18) | (0.18) | (0.19) |
| Professorship | | | 0.307* |
| | | | (0.19) |
| Share of female candidates | | -0.094 | -0.132 |
| | | (0.44) | (0.50) |
| *Variables at the candidate level* | | | |
| PhD supervisor in the committee | | | 0.381** |
| | | | (0.16) |
| PhD from the same institute | | | 0.326*** |
| | | | (0.10) |
| Standardised age-adjusted H-index | | 0.280*** | 0.286*** |
| | | (0.05) | (0.05) |
| Gender × Discipline Fixed Effects | No | No | Yes |
| Gender × University Fixed Effects | No | Yes | Yes |
| Gender × Year Fixed Effects | Yes | Yes | Yes |
| Observations | 553 | 553 | 553 |

$*p < 0.10, **p < 0.05, ***p < 0.01$. The gender variable is coded as 1 for a woman and 0 for a man. Robust standard errors in parentheses. The treatment variable is based on mean pre-reform gender parity at the field*university level. Variables interacted with the gender of the candidate should be interpreted as the differential effect of the variable for women with respect to men. Contests where only one candidate was ranked, or where only applicants of a specific gender applied are dropped from the analysis.

***Le LIEPP (Laboratoire interdisciplinaire d'évaluation des politiques publiques) est un laboratoire d'excellence (Labex). Ce projet est distingué par le jury scientifique international désigné par l'Agence nationale de la recherche (ANR). Il est financé dans le cadre des investissements d'avenir.***

*(ANR-11-LABX-0091, ANR-11-IDEX-0005-02)*

***www.sciencespo.fr/liepp***

**A propos de la publication**

**Procédure de soumission :**
Rédigé par un ou plusieurs chercheurs sur un projet en cours, le *Working paper* vise à susciter la discussion scientifique et à faire progresser la connaissance sur le sujet étudié. Il est destiné à être publié dans des revues à comité de lecture (peer review) et à ce titre répond aux exigences académiques. Les textes proposés peuvent être en français ou en anglais. En début de texte doivent figurer : les auteurs et leur affiliation institutionnelle, un résumé et des mots clefs.

Le manuscrit sera adressé à : liepp@sciencespo.fr

Les opinions exprimées dans les articles ou reproduites dans les analyses n'engagent que leurs auteurs.

**Directeur de publication :**
Bruno Palier

**Comité de rédaction :**
Carolina Alban Paredes, Andreana Khristova

Sciences Po - LIEPP
27 rue Saint Guillaume
75007 Paris - France
+33(0)1.45.49.83.61
liepp@sciencespo.fr