UNIVERSITÀ DELLA CALABRIA

Dipartimento di Economia e Statistica
Ponte Pietro Bucci, Cubo 0/C
87036 Arcavacata di Rende (Cosenza)
Italy
http://www.ecostat.unical.it/

Working Paper n. 06 - 2011

GENDER DISCRIMINATION AND EVALUATORS' GENDER: EVIDENCE FROM THE ITALIAN ACADEMY

Maria De Paola
Dipartimento di Economia e Statistica
Università della Calabria
Ponte Pietro Bucci, Cubo 1/C
Tel.: +39 0984 492459
Fax: +39 0984 492421
e-mail: m.depaola@unical.it

Vincenzo Scoppa
Dipartimento di Economia e Statistica
Università della Calabria
Ponte Pietro Bucci, Cubo 1/C
Tel.: +39 0984 492464
Fax: +39 0984 492421
e-mail: v.scoppa@unical.it

Giugno 2011

Gender Discrimination and Evaluators' Gender:

Evidence from the Italian Academy

Maria De Paola, Vincenzo Scoppa*

Department of Economics and Statistics, University of Calabria

Abstract: Relying on a natural experiment consisting in 130 competitions for promotion to associate and full professor in the Italian University, we analyze whether gender discrimination is affected by the gender of evaluators. Taking advantage of the random assignment of evaluators to each competition, we examine the probability of success of each candidate in relation to the committee gender composition, controlling for candidates' scientific productivity and a number of individual characteristics. We find that female candidates are less likely to be promoted when the committee is composed exclusively by males, while the gender gap disappears when the candidates are evaluated by a mixed sex committee. Results are qualitatively similar across fields and type of competitions. The analysis of candidates' decisions to withdraw from competition highlights that gender differences in preferences for competition play only a minor role in explaining gender discrimination. It also emerges that withdrawal decisions are not affected by the committee gender composition and therefore the gender discrimination is not related to self-fulfilling expectations.

JEL classification: D72, D78; J45, J71

Keywords: Gender Discrimination; Evaluators' Gender; Affirmative Actions; Academic Promotion ; Withdrawal Decision; Natural Experiment; Random Assignment.

1. Introduction

Female education levels have recently improved considerably and female labor force participation has risen in most countries. However, in many spheres of human life gender inequality is still pervasive. A huge literature shows that female employees earn less than males even when they have the same levels of education, work experience and professional qualification (see, among others, Blau and Kahn, 2003; Altonji and Blank, 1999; Weichselbaumer and Winter-Ebmer, 2005). A number of papers show that in many countries the gender wage gap is increasing across the wage distribution (Arulampalam *et al.*, 2007; Albrecht *et al.*, 2003) and that women face the so called "glass ceiling", that is, they remain greatly underrepresented in higher paying jobs and in top positions, probably because promotion procedures favor men rather than women. For example, some recent works examining promotions and pay in the academic

* Department of Economics and Statistics, University of Calabria, 87036 Arcavacata di Rende (CS), Italy. E-mail: m.depaola@unical.it; v.scoppa@unical.it. We would like to thank Davide Fiaschi, Anna Giunta, Luca Gori, Michela Ponzo and Manuela Stranges.

labor market show that women suffer a disadvantage in promotions and a within-rank pay gap (Blackaby *et al.*, 2005; McDowell *et al.*, 1999; Ginther and Kahn, 2004).

An interesting issue is to what extent discrimination depends on the gender of evaluators. At the best of our knowledge only a few works have tried to examine this issue. Two recent papers by Bagues and Esteve-Volart (2010) and by Zinovyeva and Bagues (2011), based, respectively, on a recruitment procedure for positions in the Spanish Judiciary and on competitions to associate and full professor positions in Spain, reach rather ambiguous results. Whereas from Bagues and Esteve-Volart (2010) it emerges that female candidates are less likely to be hired when the randomly assigned selection committee is characterized by a higher percentage of female evaluators, Zinovyeva and Bagues (2011) show that committees with a relatively larger share of females reduce gender discrimination against women in competitions to full professors positions, but they find no statistically significant effect as regards competitions to associate professor.

A different approach has been taken by some works investigating the impact produced by female boss on female employees. Bell (2005) using US data finds that women-led firms hire more top executive women and pay higher wages to female workers compared to men-led firms. Similarly, Cardoso and Winter-Ebmer (2007), using data from Portugal, show that female leadership in firms leads to higher wages for women and lower wages for males. Goldin and Rouse (2000) show that female musicians have increased their probability of being hired in prevailingly male symphony orchestra after the adoption of "blind" auditions with a "screen" to conceal the candidate's identity from the jury. In an educational context, Lavy (2008), analyzing the existence of gender discrimination in teachers' evaluation of students comparing results in a blind and in a non-blind test, shows that the gender bias is sensitive to the gender of evaluators, but the direction of the effect varies across disciplines.

In this paper we try to shed more light on this issue providing new evidence on whether the gender of evaluators matters for discrimination. We base our analysis on a natural experiment involving the Italian academic promotion system for associate and full professor positions. Our framework shares with the papers of Bagues and Esteve-Volart (2010) and Zinovyeva and Bagues (2011) the same identification strategy, based on the random assignment of evaluators to competitions.

Thanks to the random assignment procedure followed in Italy to select the members of evaluation committees for competitions to associate and full professor positions opened in 2008, we are able to estimate the probability of success of candidates in relation to the committee gender composition, avoiding endogeneity problems deriving from unobservable factors that may be correlated with committees' and candidates' characteristics.

Unfortunately, data on these competitions are not readily available and have to be collected reading the official reports produced by each committee. Since collecting data on all promotion procedures would have been an unmanageable task, we have decided to focus on promotion procedures in only two fields: Economics and Chemistry. More precisely, we use data on 130 public competitions involving about 1,000 candidates evaluated by 650 professors.

For each committee member and for each candidate we have collected data on the number of publications, the number of citations, *h* and *g* indexes, and on the university where they worked at the time of the competition. We have used these information to build indicators of candidates' and committees' scientific productivity, to identify whether candidates are insiders in the university opening the vacancy and to find out professional networks between candidates and committee members.

Controlling for all these factors, we have estimated the probability of success of each candidate. From our analysis it emerges that female candidates have a lower probability of success compared to their male counterparts (3.7 percentage points less). Since we control for a large number of individual characteristics and for a number of quite reliable measures of individual productivity, we are confident that – in comparison to the large part of the literature on gender wage gap (see Cahuc and Zylberberg, 2004) – our results are less affected by problems deriving from unobservable characteristics, unbalanced across gender, that may determine individual earnings.

As regards the effect of the committee gender composition, we find that female candidates are significantly less likely to be promoted when the randomly assigned committee is composed exclusively by males: in this case the probability of success of females is about 6-7 percentage points less. On the contrary, the presence of female members in the committee allows to overcome almost completely discrimination against women. The impact of committee gender composition turns out to be about the same in specifications controlling for individual fixed effects preventing any bias arising from correlation of gender with unobservable characteristics. This result holds true both for the Economics and Chemistry fields. As regards heterogeneous effects across different type of positions, we find that in competitions to associate professor, committees composed exclusively by males operate a stronger discrimination against women than that emerging in competitions to full professor positions. Moreover, the improvement in female outcomes produced by a mixed sex committee is smaller in magnitude in competitions to associate professor.

One new and interesting characteristic of our data is that we observe whether applying candidates decided to withdraw from competition during the evaluation process. This allows us to investigate whether the observed gender discrimination is related to the fact that women tend to shy away from competition and to be less career oriented (Booth, 2009; Manning and Saidi, 2010; Bertrand, 2011). We find that females are more likely to withdraw from competitions to associate professor, while there are no statistically significant differences in competitions to full professor positions. Nevertheless, for both types of competitions, we show that gender discrimination emerges also when we consider exclusively the sample of subjects that have not withdrawn from competition, implying that gender differences in preferences for competition play only a minor role in explaining the bias against women emerging in our analysis.

The availability of data on withdrawal decisions allows us also to evaluate whether the effect of committee gender composition on female probability of success depends on self-fulfilling expectations, since, at least in principle, women may decide to retire their candidacy once they know the committee

composition. We do not find any statistically significant effect of the committee gender composition on female's probability of withdrawing from competition. Moreover, the positive impact of mixed sex committees on female candidates' probability of success persists also when we exclude from our sample those who have withdrawn from competition.

The paper is organized as follows. Section 2 presents the Italian academic promotion system and describes the data used in our analysis. In section 3 we carry out some random assignment checks. In section 4 we show our estimation results on the impact of committee gender composition on female candidates' probability of success. Section 5 is devoted at investigating differences across types of position and fields. In section 6 we investigate whether our results are robust to the exclusion of those candidates that have decided to withdraw from competition and whether the probability of withdrawing is related to the gender composition of the committee. In Section 7 we present candidate fixed effects estimates. In section 8 we offer a set of robustness checks. Section 9 concludes.

2. Institutional Background and Data

Italy is one of the worst performing countries in terms of gender equality: in 2009, the Gender Gap Index ranks Italy at the 72nd position, below Kazakhstan and Ghana. Women are underrepresented both in the public and in the private sector. Only 20% of seats in the parliament are held by women and only 3% of the 50 largest companies' board directors are women. As far as the academia is concerned, women account for 45% of assistant professors, 34% of associate professors and for 20% of full professors. Although the number of women in the lower ranks has grown over time, the increase has been modest among higher positions.

The rules governing careers in the Italian Universities have changed over time. Abandoning a centralized and nationwide competition, a new mechanism was implemented for promotion to associate and full professor positions since 1999: each university willing to fill a vacancy initiated a competition and a committee of five members was selected to choose two or three winners (so called "idonei"). One member of the committee was appointed by the university opening the vacancy and the remaining four were elected by all professors in the field.

These rules were strongly criticized because elected committee members were not typically chosen with the aim to screen the best candidates but according to agreements among influential members of the academia, with the result that promotions were far from being related to candidates' scientific productivity.[1] Nevertheless, in 2008, under this system, a huge number of vacancies (695 positions for full professors and 1,110 positions for associate professors) were opened by Italian Universities. At the end of

[1] Analysing the working of the Italian academic competitions, Perotti (2002) describes the system as follows: "University X wants to promote its own insider, and initiates a competition. The commissioner from university Y supports "idoneità" [promotion] for the insider of university X, with the mutual understanding that university X will return the favour in the future when it comes to promoting university Y's insider".

2008, the Italian Government, worried of the outcomes that could arise by the system in force, has decided to change the rules governing promotions to associate and full professor positions. The main change has concerned the way in which committees are selected: it has been established that four members out of five have to be randomly selected (among all the full professors in each field)[2] instead of being elected, while, as in the previous system, one member is appointed by the university opening the vacancy.

Committee members meet to evaluate candidates and at the end of the evaluation process two winners for each evaluation procedure are selected. While in competitions to full professor candidates are evaluated exclusively on the basis of their CV, in competitions to associate professor skills shown by candidates in a teaching lecture are also taken into account. As in the previous system, the University that has initiated the competition can decide to appoint one of the winning candidates as professor, while the other can be appointed by another university within three years.

As explained above, data on competitions have to be collected reading the final report produced by each committee at the end of the evaluation process. Due to the huge amount of work related to data collection, we have chosen to focus our attention exclusively on competitions undertaken in two relatively large fields: Economics (5 sub-fields)[3] and Chemistry (10 sub-fields).[4] We have chosen these two fields with the aim of analyzing both a scientific and a social science field. Among scientific fields, Chemistry was characterized by a quite large proportion of females, while other fields, such as Physics or Engineering were excluded due to the extremely small number of female evaluators (mirroring the scarce presence of females in the field). Among social science fields, we have focused on economics because it was easier to find measures of individual productivity compared for example to Humanities or Sociology.

By February 2011, 52 competitions (31 to associate professor and 21 to full professor) were concluded in the Economics field, while in Chemistry 78 competitions (46 to associate professor and 32 to full professor) were completed. As a consequence, we end up with 130 evaluation procedures, involving 1,007 candidates and 650 committee members. The average number of competitors for each competition is equal to 17.53. Candidates were allowed to apply to a maximum of 5 different competitions. Each candidate has applied on average to 2 competitions. The total number of observations at the candidate-competition level is equal to 2279.

During the evaluation process about 27% of candidates decided to withdraw from competition. Withdrawals are more frequent in competitions to associate professor positions (43.6%) than in competition to full professor positions (7.8%). The sample including only the candidates that maintain their candidacy until the conclusion of the evaluation procedure is made of 1,652 observations.

[2] The selection is carried out by the officials of the Ministry of Education, University and Research, through a computerized random procedure certified by a notary.

[3] For example, sub-fields in economics are Econometrics, Public Economics, Applied Economics etc., while for Chemistry sub-fields are Organic Chemistry, Inorganic Chemistry, Physical Chemistry etc.

[4] In Economics, 28% of professors are females (women account for 42% of assistant professors, 26% of associate professors and 16% of full professors). In Chemistry, 42% of professors are females (women account for 57% of assistant professors, 40% of associate professors and 18% of full professors).

We have collected the list of evaluators, candidates and winners from the final reports produced by each committee. The gender has been inferred from the first name. Age has been taken from official reports or searching CVs on-line. In the few cases in which we were not able to find the year of birth we have imputed it as the year of graduation minus 24 (the age at which typically high ability students graduate).

To gather information on the scientific productivity of candidates and evaluators we have used the "Publish or Perish" software based on Google Scholar. More precisely, we have collected data on the number of publications, citations, *h* and *g* indexes, for each individual at the data of conclusion of each competition. We have decided to consider the publications until this date instead of until the date of application since long delays typically occur from when papers are accepted for publication (and candidates include them in their CVs) and when publications appear as published in scientific journals.

Using data on the number of publications and citations and on the *h* and *g* indexes, we have undertaken a principal component analysis to obtain a comprehensive measure of individual productivity (only the first component was considered), which we call *Productivity*. For each candidate we build *Relative Productivity* as the difference between his/her *Productivity* minus the average productivity of the other candidates in the competition. Moreover, this measure of productivity is used to calculate for each evaluation committee the average productivity of evaluators, considering only the four randomly selected members.

The affiliations of both evaluators and candidates have been obtained from the Ministry of Education, University and Research (MIUR)[5] and used to build a dummy variable *Insider* taking the value of one for candidates who work in the university opening the vacancy. Moreover, we build an indicator of professional networks between candidates and committee members, *Connections*, taking the value of one when there is at least a committee member (excluding the internal appointed evaluator) from the same university as the candidate and zero otherwise.

Descriptive statistics for candidates and for evaluators are reported in Table 1. The percentage of female candidates is about 40%, higher in competitions to associate professors (45%) than in competitions to full professors (33%) (35% in Economics and 43% in Chemistry). Candidates to full professor positions over their lifetime have published on average 61 works receiving 469 citations, whereas the average number of publications of candidates to associate professor was 41 with 274 citations. About 15% of candidates are insiders and 10% of them has connections with at least one member of the committee. The great majority of candidates is performing an academic job (90%). On average, candidates are 44.7 years old, candidates to associate professor positions are typically younger (41.7) than candidates to full professor positions (48.3) and chemistry candidates are older (46.7) than economics ones (42.1).

As regards evaluators, we focus our attention exclusively on the four randomly selected committee members and neglect the internal commissioner since the individual characteristics of the latter could be correlated to unobservable determinants of success of candidates. About 16% of the randomly

[5] From the web page: http://cercauniversita.cineca.it/php5/docenti/cerca.php

selected evaluators are females. 55% of committees are composed exclusively by males, 31% has one female member, and 10% and 4% percent of committees has, respectively, 2 and 3 female members. Given this distribution, we build a dummy variable *Females in Committee* taking the value of one when at least one female was among the committee members: 44.6% of committees have among their members a female evaluator. The average age of evaluators is 60 and about 28% of evaluators are from Universities of the South of Italy. On average committee members over their lifetime have published 82 papers receiving 779 citations.

Table 1. Descriptive Statistics

	Mean	*St. Dev*	*Min*	*Max*	*Observations*
		Candidates			
Success	0.113	0.317	0	1	2279
Associate Professor	0.122	0.327	0	1	1024
Full Professor	0.103	0.304	0	1	1255
Female	0.397	0.489	0	1	2279
Associate Professor	0.453	0.498	0	1	1024
Full Professor	0.328	0.469	0	1	1255
Number of Papers	50.491	37.740	0	269	2279
Associate Professor	41.909	30.080	0	199	1024
Full Professor	61.010	43.144	0	269	1255
Citations	362.502	491.504	0	4485	2279
Associate Professor	274.811	384.017	0	4485	1024
Full Professor	469.976	579.806	0	4431	1255
h-index	8.812	5.404	0	36	2279
Associate Professor	7.726	4.814	0	36	1024
Full Professor	10.144	5.779	0	35	1255
g-index	14.646	9.409	0	65	2279
Associate Professor	12.778	8.348	0	64	1255
Full Professor	16.935	10.109	0	65	1024
Relative Productivity	0	1.758	-5.256	10.593	2279
Associate Professor	0	2.095	-5.256	9.387	1024
Full Professor	0	1.426	-5.081	10.593	1255
Insider	0.147	0.355	0	1	2279
Associate Professor	0.168	0.329	0	1	1024
Full Professor	0.123	0.374	0	1	1255
Connections	0.103	0.305	0	1	2279
Associate Professor	0.112	0.317	0	1	1024
Full Professor	0.091	0.289	0	1	1255
Age	44.697	6.771	29	69	2279
Associate Professor	41.728	5.330	30	69	1024
Full Professor	48.337	6.572	29	67	1255
University Job	0.902	0.298	0	1	2279
Associate Professor	0.873	0.333	0	1	1024
Full Professor	0.936	0.244	0	1	1255
Withdrawn	0.275	0.446	0	1	2279
Associate Professor	0.436	0.496	0	1	1024
Full Professor	0.078	0.268	0	1	1255
		Committees' members			
Females in Committee	0.446	0.4990	0	1	130
% Females in Committee	0.156	0.2047	0	0.75	130
Age Com. members	60.480	3.5786	49.5	69	130
Perc. of Com. members South	0.283	0.2459	0	1	130
Number of papers Com. members	82.943	88.7533	1	819	520
Number citations Com. members	779.916	1591.934	0	22370	520
Associate Position	0.592	0.4933	0	1	130
Economics	0.400	0.4918	0	1	130

3. Random Assignment Checks

Our identification strategy is based on the random assignment of committee members to each competition. To investigate the randomness of the assignment mechanism, we regress a number of individual characteristics of candidates participating at each competition (the percentage of female candidates, the average productivity of candidates, the number of candidates competing for the position, the percentage of insider candidates) on the dummy *Females in Committee*, controlling for sub-field dummies (since the random assignment of evaluators to competitions was conditional on sub-fields) and for the type of position (a dummy taking value of one for competitions to associate positions and zero otherwise). Moreover, since in building *Females in Committee* we only consider the percentage of females among the randomly selected evaluators without considering the committee member appointed by the university opening the vacancy, we also check for any possible correlation between the characteristics of the internal member and the presence of females among the randomly selected committee members.

The estimation results are reported in Table 2. In column (1) we show that the correlation between the percentage of female candidates and the presence of females evaluators is far from being statistically significant. Similarly, there is no statistically significant correlation between the presence of females in the committee and the average productivity of candidates (column 2), the number of competing candidates (column 3) and the percentage of internal candidates (column 4). In columns 5 and 6 we show that the presence of female evaluators in the committee is not related to the gender or to the scientific productivity of the internal commissioner.

We also checked whether the predetermined characteristics are related to the percentage of females in committees (*% Females in Committee*) instead of using the dummy *Females in Committee*, obtaining very similar results (not reported).

As an alternative check, we have used our measures of committee gender composition as dependent variables, regressing them on the full set of variables describing predetermined characteristics at the competition level (percentage of female candidates, average productivity, number of competitors, percentage of insiders) and have tested for the joint significance of these covariates. Results (not reported) do not allow us to reject the null hypothesis of zero effects (the F-test is equal to 0.67 with a p-value of 0.68).

Since female committee members were not sorted into competitions according to the characteristics of the candidates or of the internal commissioner, we conclude that the assignment of evaluators to each competition has been effectively random.

Table 2. Regressions for Random Assignment Checks

	(1) % Female Candidates	(2) Candidates' Average Productivity	(3) Number of Competitors	(4) % Insiders	(5) Female Internal Commissioner	(6) Productivity Internal Commissioner
Females in Committee	0.002	0.064	0.804	0.014	0.004	-0.428
	(0.026)	(0.114)	(1.262)	(0.023)	(0.060)	(0.381)
Observations	130	130	130	130	130	130

Data at competition level. The dependent variable is reported at the top of each column. Standard errors are reported in parenthesis. In all the regressions we control for sub-field dummies and for type of position dummy.

4. Gender Discrimination and Evaluators' Gender: The Empirical Findings

To uncover the effect of committee gender composition on the probability of success of candidates we estimate the following model:

[1] $$Success_{ij} = \beta_0 + \beta_1 Female_i + \beta_2 Females\ in\ Committee_j + \beta_3 Female_i * Females\ in\ Committee_j + \phi X_{ij} + \mu_j + \lambda_j + \varepsilon_{ij}$$

where the dummy variable $Success_{ij}$, taking value of 1 if candidate *i* has won competition *j*, depends on the candidate gender $Female_i$, on a vector X_{ij} of the candidate characteristics (including scientific productivity and the dummies $Insider_{ij}$ and $Connections_{ij}$), our indicator of committee gender composition, $Females\ in\ Committee_j$, a dummy for the type of position, λ_j, and dummies for scientific sub-fields μ_j. To investigate whether the probability of success of candidates is affected by the gender composition of the committee, we include among our regressors the interaction term between $Female_i$ and $Females\ in\ Committee_j$. Therefore, the coefficient β_1 measures the effect of being a female on the probability of success when the evaluation committee is composed exclusively by men, while $\beta_1 + \beta_3$ represents the extent of female discrimination (if any) when there is at least a female among the committee members.

Equation [1] is estimated using a probit model. Marginal effects are reported in Table 3. We base our analysis on the whole sample of candidates applying for a position to associate or full professor, independently from their effective participation in the competition (in Section 6 we exclude candidates withdrawing from competition). In all the regressions standard errors are clustered at the competition level to take into account that common shocks may affect the performance of all candidates participating to the competition.

In the first specification of Table 3 (column 1) we estimate the difference in the probability of success between males and females controlling for scientific sub-field dummies, type of position and number of competing candidates, without taking into account other candidates' characteristics: a female has a lower probability of success of about 4.7 percentage points (significant at the 1 percent level). Since

on average the probability of success for a male is about 13%, females suffer a reduction of about 36% in the chances of winning a competition.

In column 2, in order to avoid the bias that may derive from the fact that candidate's gender may be related to some individual features affecting the probability of success, we include among regressors the comprehensive measure of scientific productivity, *Relative Productivity*, and the dummy variables *Insider*, *Connections* and *University Job*. The candidate's scientific productivity contributes to the probability of winning though the effect is quite small in magnitude: an increase of one standard deviation in *Relative Productivity* produces an increase in the probability of success of about 3.5 percentage points. On the other hand, it emerges that being an insider strongly improves the probability of success (by 28 percentage points). *Connections* are also relevant and increase the probability of success by 7 percentage points. The dummy *University Job* is not statistically significant. Importantly, also controlling for these characteristics, it emerges that females suffer a reduction of 3.7 percentage points in the probability of success.

In column 3 we estimate specification 1 adding as regressors the dummy $Females in Committee_j$ and the interaction term $Female_i * Females\ in\ Committee_j$. It emerges that the presence of female members in the committee increases the probability of success of female candidates. More precisely, all-males committees reduce the probability of success of female candidates by 7.6 percentage points, while mixed sex committees eliminates the gender discrimination against women. When among the committee members is present at least a female, the difference between males and females in the probability of success turns out to be equal to -0.014 with a *p*-value of 0.489.[6]

To analyse whether the gender discrimination implemented by all-males committees is related to individual characteristics, in column 4 we include the controls considered in column 2. Our results remain substantially unchanged as we find that committees composed exclusively by men discriminate against women, reducing their probability of success by about 6.4 percentage points. However, gender discrimination disappears when candidates are judged by a mixed sex committee: the difference between males and females is equal to -0.007 (*p*-value =0.737).

Since we are controlling for scientific productivity, the estimated gender difference cannot be imputed to any difference of productivity between males and females. Moreover, given the controls for *Insider* and *Connections*, we are able to exclude that the uncovered effect depends on possible differences between males and females in the probability of being an insider or having a connection.

To have an idea of the magnitude of the effect produced by a mixed sex committee on female candidates' probability of success, consider that it is equivalent to the improvement deriving from an increase of 2 standard deviations in a candidate's *Relative Productivity*.

[6] As in probit models the interaction effect is the cross-partial derivative of the expected value of the dependent variable, the interaction effect cannot be interpreted straightforwardly. To investigate the effect of mixed sex committee on female candidates' probability of success we have used the Stata command *predictnl*.

These results hold true also when – instead of using *Relative Productivity* – we consider separately our different measures of individual scientific productivity. In column 5 we report results obtained using the relative *h* index. Similar findings are obtained also using alternatively the number of publications, citations or the *g* index (not reported).

In column 6, instead of the dummy variable $Females\ in\ Committee_j$, we consider the *% Females in* $Committee_j$ among the randomly selected members of the committee as a measure of committee sex composition. From the coefficient on the interaction (*Female*)*(*% Females in* $Committee_j$) it turns out that one more female in the committee increases the probability of success of female candidates of 2.2 percentage points. The presence of two female members in the committee reduces to zero the bias against women.

Table 3. Estimates of the Probability of Success

	(1)	(2)	(3)	(4)	(5)	(6)
Female	-0.047***	-0.037***	-0.076***	-0.064***	-0.065***	-0.051***
	(0.014)	(0.012)	(0.017)	(0.016)	(0.016)	(0.015)
Female*(Females in Com.)			0.076**	0.072*	0.073*	
			(0.037)	(0.037)	(0.036)	
Female*(% Females in Com.)						0.089
						(0.063)
Females in Com.			-0.021**	-0.020*	-0.020**	
			(0.010)	(0.010)	(0.010)	
% Females in Com.						-0.023
						(0.026)
Relative Productivity		0.020***		0.020***		0.020***
		(0.003)		(0.003)		(0.003)
Insider		0.283***		0.283***	0.282***	0.283***
		(0.029)		(0.029)	(0.029)	(0.029)
Connections		0.070***		0.071***	0.069***	0.070***
		(0.026)		(0.026)	(0.026)	(0.026)
University Job		-0.007		-0.006	-0.011	-0.007
		(0.019)		(0.020)	(0.020)	(0.020)
Age		-0.001		-0.001	-0.001	-0.001
		(0.001)		(0.001)	(0.001)	(0.001)
Relative h-index					0.008***	
					(0.001)	
Observations	2279	2279	2279	2279	2279	2279
Pseudo R-squared	0.041	0.166	0.066	0.170	0.167	0.167

Notes: The Table reports marginal effects of Probit estimates (evaluated at the mean values of the explanatory variables in the sample). The dependent variable is *Success*. In all regressions we control for sub-field dummies, type of position dummy and the number of candidates. Standard errors (corrected for heteroskedasticity and clustered at the competition level) are reported in parentheses. The symbols ***, **, * indicate that coefficients are statistically significant, respectively, at the 1, 5, and 10 percent level.

In Figure 1 we plot the probability of success of male and female candidates, separately for all-males and mixed sex committees, in relation to their relative scientific productivity based on the results of column 4 (Table 3). The vertical distance between the continuous line (above in the figure) and the dashed line (below in the figure) represents the gender discrimination when the judging committee is composed exclusively by males. As shown in the Figure, this gap tends to close when the judging committee is composed also by female members: the two lines representing male's (short dashed) and female's (dotted) probability of success with a mixed sex committee are very close to each other.

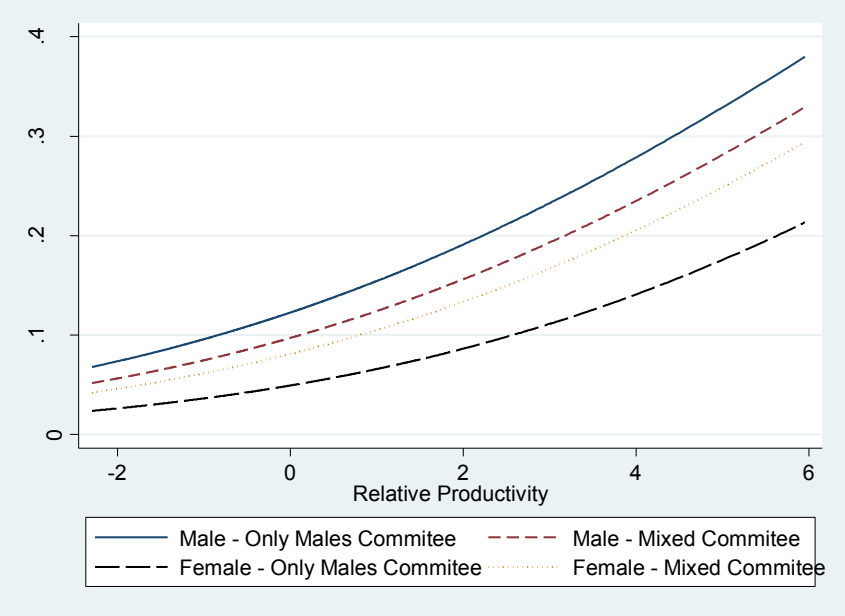

Figure 1. Probability of success of male and female candidates according to the gender composition of the committee

Results similar to those shown in Table 3 are also obtained using a linear probability model (results not reported) instead of a probit. We have also experimented using multi-way clustering for standard errors, as suggested by Cameron *et al.* (2006), at competition and candidate level. The significance of the coefficients of interest does not change and, if anything, it slightly improves.

5. Are the Effects of Mixed Sex Committee Heterogeneous across Fields and Positions?

In this section we investigate whether the effects of a mixed sex committee are heterogeneous across fields and in relation to the type of position for which promotion is decided. We re-estimate specification 4 of Table 3 separately for competitions to associate and to full professor positions (respectively column 1 and column 2 of Table 4). It emerges that in competitions to associate professor committees composed exclusively by males operate a stronger discrimination against women in comparison to that emerging in competitions to full professor positions. Moreover, the improvement produced by a mixed sex committee is smaller in magnitude in competitions to associate professor. More precisely, in the latter competitions, when evaluators are exclusively males, females experiment a reduction in the probability of success of 8 percentage points, while the presence of a mixed sex committee reduces the bias against women to 4.7 percentage points (statistically significant at the 5 percent level). On the other hand, in competitions to full professor, when evaluated by an all-males committee, females' candidates face a reduction in the probability of success of 5 percentage points. In this type of competition, the bias against women vanishes when the evaluation committee is composed also by female members: the difference in the probability of success between males and females is 0.031 with a *p*-value of 0.280.

All in all, these results suggest that in both type of competitions the presence of female evaluators enhances the probability of success of female candidates and helps at reducing the bias against women produced by all-males committees.

In columns 3 and 4 of Table 4 are reported estimation results for competitions taking place in the fields of, respectively, Economics and Chemistry. All-male committees are gender biased in both fields, but the bias is larger in competitions taking place within the Chemistry field: females experiment a reduction in the probability of success of 7 percentage points in Chemistry and 5 percentage points in Economics. The effect of a mixed sex committee goes in the same direction in both fields and allows male and female candidates to face equality of treatment.

Table 4. Heterogeneous Effects across Positions and Fields

	(1) Associate Professor	(2) Full Professor	(3) Economics	(4) Chemistry
Female	-0.080***	-0.050*	-0.051**	-0.070***
	(0.024)	(0.021)	(0.021)	(0.023)
Female*(Females in Com.)	0.038	0.120	0.063	0.070
	(0.041)	(0.067)	(0.050)	(0.048)
Females in Com.	-0.007	-0.028**	-0.020	-0.020
	(0.015)	(0.014)	(0.013)	(0.015)
Relative Productivity	0.024***	0.015***	0.024***	0.017***
	(0.005)	(0.003)	(0.004)	(0.004)
Insider	0.292***	0.263***	0.222***	0.329***
	(0.038)	(0.051)	(0.046)	(0.037)
Connections	0.065**	0.067**	0.112***	0.049
	(0.034)	(0.038)	(0.042)	(0.034)
University Job	0.018	-0.047	0.011	-0.017
	(0.023)	(0.037)	(0.022)	(0.035)
Age	0.003**	-0.004**	-0.001	-0.001
	(0.001)	(0.001)	(0.002)	(0.001)
Observations	1235	991	997	1229
Pseudo R-squared	0.182	0.181	0.154	0.188

Notes: The Table reports marginal effects of Probit estimates (evaluated at the mean values of the explanatory variables in the sample). The dependent variable is *Success*. In all regressions we control for sub-field dummies, type of position dummy and the number of candidates. Standard errors (corrected for heteroskedasticity and clustered at the competition level) are reported in parentheses. The symbols ***, **, * indicate that coefficients are statistically significant, respectively, at the 1, 5, and 10 percent level.

Since we have found quite similar results across different type of positions and different fields, it seems hard to argue that the discrimination of male evaluators against female candidates is related to unobserved quality or to the fact that male candidates perform better in some unobserved "task" that are particularly appreciated by males. In fact, we would have expected unobserved quality to play a major role in the fields in which the use of biblio-metric indexes (as those used in our analysis) to judge candidate's quality is less diffuse, or in competitions to associate professor, where candidates are evaluated also in relation to their teaching skills (which we do not observe). However, from our estimates it emerges that the gender discrimination practiced by committees composed exclusively by males is stronger in Chemistry than in Economics, in spite of the fact that the indexes we use to measure individual productivity are more widely used in Chemistry than in Economics. Furthermore, we obtain very similar results for competitions to associate and to full professor positions even if unobserved teaching abilities should matter more in competition to associate professor.

6. Discrimination, Preferences for Competition and Self-fulfilling Expectations

In principle, males' and females' expectations might differ ex ante and differences in expectations may affect outcomes ex post. More specifically, female candidates might believe that they are less likely to be promoted when the evaluating committee is composed exclusively by men and, as a consequence, once informed about the gender composition of the committee, they may decide to withdraw from competition. In this case, the outcome we observe would be driven by women's expectations rather than be determined by the effective behavior of all-male committees.

Thanks to the availability of information on the behavior of each candidate, we are able to check whether our results are driven by the fact that female candidates retire once they knew the gender composition of the committee. At this aim, we have excluded from our sample all the candidates who have withdrawn from competition (about 27% of candidates).

In Table 5 we present the first four specifications reported in Table 3, plus the results obtained separately for competitions to associate and full professor positions, considering only the sample of candidates that have maintained their candidacy until the end of the competition process. Our previous findings are confirmed. Again, it emerges that committees composed exclusively by males discriminate against women reducing their probability of success: the effect is even higher in magnitude than that emerging from previous estimates based on the full sample of applicants. On the other hand, mixed sex committees tend to reduce gender discrimination. Therefore, our results are not driven by self-fulfilling expectations.

Results similar to those reported in columns 1 and 2 of Table 4 are found considering separately competitions to associate and full professor: discrimination against women is stronger in competitions to associate professor, where we also observe a smaller positive impact of mixed sex committees compared to that emerging in competitions to full professor.

These findings also imply that differences between males and females in preferences for competition (see Bertrand, 2011) play a minor role in explaining female worse outcomes. In fact, in the estimates of Table 5 we have considered subjects who have applied for a position and who have maintained their candidacy until the conclusion of the evaluation process showing in this way their willingness to compete.

Table 5. Estimates of the Probability of Success Excluding Withdrawals

	(1) All	(2) All	(3) All	(4) All	(5) Associate Professor	(6) Full Professor
Female	-0.063***	-0.050***	-0.098***	-0.085***	-0.143***	-0.056***
	(0.018)	(0.017)	(0.022)	(0.022)	(0.047)	(0.022)
Female*(Females in Com.)			0.087**	0.092*	0.063	0.115**
			(0.047)	(0.048)	(0.074)	(0.067)
Females in Com.			-0.027**	-0.025*	-0.006	-0.029**
			(0.014)	(0.014)	(0.029)	(0.014)
Relative Productivity		0.026***		0.026***	0.043***	0.016***
		(0.004)		(0.004)	(0.013)	(0.003)
Insider		0.282***		0.283***	0.304***	0.264***
		(0.034)		(0.034)	(0.050)	(0.052)
Connections		0.085***		0.085***	0.074	0.078***
		(0.033)		(0.034)	(0.054)	(0.041)
University Job		-0.035		-0.034	-0.039	-0.100
		(0.035)		(0.035)	(0.049)	(0.052)
Age		-0.002		-0.002	0.006***	-0.005
		(0.002)		(0.002)	(0.003)	(0.002)
Observations	1652	1652	1652	1656	708	944
Pseudo R-squared	0.079	0.183	0.082	0.187	0.171	0.192

Notes: The Table reports marginal effects of Probit estimates (evaluated at the mean values of the explanatory variables in the sample). The dependent variable is *Success*. In all regressions we control for sub-field dummies, type of position dummy and the number of candidates. Standard errors (corrected for heteroskedasticity and clustered at the competition level) are reported in parentheses. The symbols ***, **, * indicate that coefficients are statistically significant, respectively, at the 1, 5, and 10 percent level.

To better understand candidates' decisions of withdrawing from competition we have also estimated a probit model considering as dependent variable a dummy taking value of 1 for candidates deciding to withdraw from competition and zero otherwise. Withdrawals are more frequent in competitions to associate professor positions since participation costs are higher due to the fact that candidates are evaluated not only in relation to their CVs but also considering their performance in a teaching lecture, typically given in the place where is located the university posting the vacancy. To take into account this aspect, we have added to the controls used in previous estimates the dummy variable *Distance* taking the value of one when the university in which the candidate is currently employed is located in a geographical area that is different from that of the university initiating the competition. We exclude for each competition the candidates that have been already promoted in some concluded competition.

In Table 6 are reported estimation results. In column 1 we estimate the difference in the probability of withdrawal between males and female controlling for scientific sub-field dummies, type of position, number of competing candidates and *Distance*, without taking into account other candidates' characteristics. It emerges that females are more likely to withdraw from competition than their male counterparts (+3.4 percentage points). The same result holds true when we add among controls *Relative Productivity*, *Insider* and *Connections* (column 2).

In column 3 we investigate whether the probability of withdrawal is affected by the committee's gender composition. We do not find any statistically significant effect.

In columns 4 and 5 we run separate regressions respectively for competitions to associate and to full professor positions. It emerges that females are more likely to withdraw from competitions to associate professor, while it does not emerge any statistically significant difference between males' and

females' withdrawal decisions as regards competitions to full professor positions. Moreover, while the gender composition of the committee does not affect withdrawal decisions in competitions to associate professors, we find that in competitions to full professor positions females are less likely to retire their candidacy when the evaluation committee is composed also by female members.

Table 6. The Determinants of Withdrawal from Competition

	(1) All	(2) All	(3) All	(4) Associate Professor	(5) Full Professor
Female	0.034*	0.036*	0.044	0.091**	0.003
	(0.019)	(0.019)	(0.027)	(0.046)	(0.014)
Female*(Females in Com.)			-0.016	-0.012	-0.037*
			(0.038)	(0.065)	(0.012)
Females in Com.			-0.013	0.013	-0.010
			(0.033)	(0.053)	(0.018)
Relative Productivity		-0.009	-0.009	-0.036***	-0.004
		(0.007)	(0.007)	(0.012)	(0.015)
Distance	0.128***	0.049**	0.050**	0.105**	0.045
	(0.019)	(0.023)	(0.023)	(0.041)	(0.031)
Associate Professor	0.327***	0.416***	0.413***		
	(0.026)	(0.034)	(0.036)		
Insider		-0.131***	-0.131***	-0.247***	-0.018
		(0.023)	(0.023)	(0.043)	(0.016)
Connections		-0.068***	-0.068***	-0.131***	-0.012
		(0.022)	(0.022)	(0.040)	(0.018)
University Job		-0.120***	-0.126***	-0.073*	-0.223***
		(0.044)	(0.043)	(0.044)	(0.076)
Age		0.000	-0.000	-0.000	-0.001
		(0.002)	(0.002)	(0.002)	(0.001)
Observations	2090	2090	2090	1146	925
Pseudo R-squared	0.206	0.227	0.229	0.113	0.251

Notes: The Table reports marginal effects of Probit estimates (evaluated at the mean values of the explanatory variables in the sample). The dependent variable is *Withdrawal*. In all regressions we control for sub-field dummies, type of position dummy and the number of candidates. Standard errors (corrected for heteroskedasticity and clustered at the competition level) are reported in parentheses. The symbols ***, **, * indicate that coefficients are statistically significant, respectively, at the 1, 5, and 10 percent level.

7. Candidate Fixed Effects Estimates

Estimates discussed in the previous sections are unbiased only if unobserved individual characteristics are uncorrelated with the explanatory variables of interest. Since evaluators were randomly assigned to competitions, this assumption should hold true. However, since a high number of candidates have participated to more than one competition (666 candidates participating on average to 2.9 competitions), in this Section we exploit the panel structure of the data and estimate equation [1] by controlling for individual fixed effects. In this way, the effect of committee gender composition is estimated exploiting the differences in outcomes for the same candidate passing from all male evaluators' to mixed sex committee's competitions. For the sake of simplicity we use a linear probability model.

In Table 7 we present the within-candidate estimates of the candidate probability of being promoted. In these estimates we lose 351 candidates who have participated to just one competition. In all specifications we include the whole set of controls for competition characteristics and for candidate's characteristics varying at competition level. In these specifications with individual fixed effects, the

coefficient on $Female_i * Females\ in\ Committee_j$ captures the change in the probability of success for a female candidate when she is evaluated by a mixed sex committee with respect to an all-male committee, while the coefficient on $Females\ in\ Committee_j$ represents the effect of a mixed sex committee on male candidates probability of success. In column 1, considering the whole sample, consistently with probit estimates shown in Table 3, it emerges that female candidates increase their probability of winning the competition when there is at least a female in the evaluation committee. On the other hand, the negative effect of a mixed sex committee on male candidates, found in probit estimates, does not appear to be statistically significant when using candidate fixed effects. This suggests that the higher chances of success encountered by females in mixed sex committees translates in worse winning prospects for the males participating to just one competition, who are typically characterized by lower productivity.

In fact, while estimates in previous sections are based on all the candidates, individual fixed effects estimates rely for each competition only on candidates participating to at least two competitions. However, the outcomes of these competitions may be affected by candidates competing just once (they may have even won) even if the latter do not contribute at defining individual fixed effects estimates.

In columns 2 and 3 are reported results considering separately competitions to associate and full professor positions. While the effect of a mixed sex committee is positive but not statistically significant for competitions to associate professor, a positive and highly statistically significant effect emerges for competitions to full professor.

Results consistent with our previous findings emerge when we run our regressions considering only candidates who have not withdrawn from competition. In column 4 are reported results considering the full sample, while in columns 5 and 6 are presented results respectively for competitions to associate and full professor positions. We find that the performance of female candidates significantly improves when there is at least a female evaluator in the judging committee. This result holds true also when we consider separately the two different type of competitions. Candidate fixed effects estimates suggest that the better treatment reserved by mixed sex committees to female candidates ends up with reducing the probability of success of males who apply to a single competition and who are probably characterized by a lower scientific productivity. Only in competitions to full professor position (see column 6) mixed sex committees produce a negative impact on the probability of success of male participating to more than one competitions.

In sum, the findings emerging from candidate fixed effects estimates are broadly consistent in terms of sign and statistical significance with the estimates of previous sections. Again, results confirm that the gender committee composition matters for discrimination against women.

Table 7. Candidates' fixed effects estimates: Linear Probability Model

	(1) All	(2) Associate Professor	(3) Full Professor	(4) All No withdrawals	(5) Ass. Prof No withdrawals	(6) Full Prof No withdrawals
Female*(Females in Com.)	0.076**	0.025	0.128***	0.114***	0.118*	0.100**
	(0.036)	(0.053)	(0.048)	(0.037)	(0.07)	(0.040)
Females in Com.	-0.014	0.019	-0.043	-0.035	0.024	-0.039*
	(0.024)	(0.040)	(0.029)	(0.023)	(0.052)	(0.023)
Observations	1938	1049	889	1351	569	782
Number of groups	666	377	302	613	325	296

Notes: The Table reports estimates obtained from a linear probability fixed effects model. The dependent variable is *Success*. In all regressions we control for candidate and competition characteristics (Relative Productivity, Insider, Connections, Number of Competitors and sub-field dummies). Standard errors (corrected for heteroskedasticity) are reported in parentheses. The symbols ***, **, * indicate that coefficients are statistically significant, respectively, at the 1, 5, and 10 percent level.

8. Robustness Checks

In this Section we present the results of a number of robustness checks. Firstly, we investigate whether our results are driven by the fact that female and male evaluators have different characteristics in terms of scientific productivity, geographical provenience and age. Secondly, we verify whether the results are affected by the characteristics of the internal member. Finally, we estimate our model excluding those fields in which the four external members of the evaluation committees have not been randomly selected among the whole body of full professors in the field, but among a number of elected full professors.

In column 1 of Table 8 we re-estimate specification 4 of Table 3 adding as controls a dummy variable, *Highly Productive Committee*, taking the value of 1 for committees with an average quality in terms of scientific productivity (measured with our comprehensive measure *Scientific Productivity*) above the mean and an interaction term between this variable and the dummy *Female*. This allow us to investigate whether female candidates are more or less favored by evaluators with different research quality. We find that the sign of the interaction coefficient is negative, but the effect in far from being statistically significant. Nevertheless, adding these controls does not change our results on the effects of the gender committee composition on discrimination against female candidates (which hold true also when we measure the research quality of both candidates and committee members using the *h* or the *g* index - not reported).

In column 2 we report estimates obtained considering the effect produced by evaluators of different age on female candidates' probability of success. As attitudes toward gender roles may change over time, it could be that older generations are more female adverse while younger ones are less likely to discriminate against women. If male evaluators are older compared to their female counterparts, it could be that the positive effect of mixed sex committees on female candidates' probability of success is not related to the gender composition of the committee but to the age of its members. To investigate this aspect we have added among our controls an interaction term between the dummy *Female* and a dummy *Above Mean Age Committee* taking the value of 1 when the average age of committee members is above the mean in the sample and zero otherwise. As shown in column 2, the age of evaluators has no effect on

the probability of success of female candidates, while the committee gender composition continues to produce the effects discussed above.

We also investigate whether evaluators working in different areas of the country show different attitudes toward women. As shown by a number of studies, women are more likely to be relegated to traditional roles in the South of Italy. As a consequence, we may expect that males working in Southern regions are more likely to be affected by gender stereotypes and to discriminate against women. At the purpose of investigating this issue, we consider the fraction of evaluators working in southern regions (% of evaluators from the South) and interact this variable with the dummy *Female*. As shown in column 3 of Table 8, female candidates are slightly more likely to suffer discrimination when the evaluators work in universities located in the South of Italy (although the *p*-value is only 0.203). However, no relevant change is observed as regards the effect of the committee gender composition on females' probability of success.

In column 4 we check the robustness of our results controlling for all the committee characteristics described above. Again our main results remain substantially unchanged.

Finally, in column 5 we have also included among controls the characteristics of the internal committee member in terms of gender and scientific productivity and have interacted these features with the dummy *Female*. The results of interest remain substantially unchanged.

Table 8. Controlling for other committee characteristics

	(1)	(2)	(3)	(4)	(5)
Female	-0.056***	-0.044*	-0.063***	0.033	-0.056***
	(0.018)	(0.023)	(0.022)	(0.029)	(0.018)
Female*(Females in Committee)	0.074**	0.064*	0.072**	0.066*	0.074**
	(0.036)	(0.036)	(0.036)	(0.036)	(0.036)
Females in Committee	-0.021**	-0.019*	-0.020**	-0.020**	-0.021**
	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)
Female*(Highly Productive Committee)	-0.021			-0.023	-0.021
	(0.023)			(0.022)	(0.023)
Highly Productive Committee	0.014			0.014	0.014
	(0.010)			(0.010)	(0.010)
Female*(Above Mean Age Committee)		-0.001		-0.005	
		(0.026)		(0.025)	
Above Mean Age Committee		-0.006		-0.004	
		(0.011)		(0.011)	
Female*(% of Evaluators from the South)			-0.065	-0.066	
			(0.051)	(0.049)	
% of Evaluators from the South			0.015	0.013	
			(0.020)	(0.020)	
Productivity Internal Member					0.001
					(0.002)
Female Internal Member					0.001
					(0.007)
Female Internal Member*Female					0.035
					(0.039)
Productivity Internal Member*Female					-0.007
					(0.005)
Observations	2279	2279	2279	2279	2279
Pseudo R-squared	0.171	0.171	0.170	0.172	0.171

Notes: The Table reports marginal effects of Probit estimates (evaluated at the mean values of the explanatory variables in the sample). The dependent variable is *Success*. In all regressions we control for sub-field dummies, type of position dummy and the number of candidates. Standard errors (corrected for heteroskedasticity and clustered at the competition level) are reported in parentheses. The symbols ***, **, * indicate that coefficients are statistically significant, respectively, at the 1, 5, and 10 percent level.

We now turn our attention to a particular feature of the Italian rules for academic promotions in force during the period under examination that may potentially bias our results. According to the rules followed for the composition of evaluation committees, in sub-fields where the number of opened vacancies was small compared to the number of available evaluators, committee members had to be randomly selected among a number of professors elected by the professors in the sub-field. The number of elected members had to be the triple of the number of evaluators needed in the *n* competitions opened in the sub-field.[7] We are confident that this institutional feature does not affect our results, since the possibility of influencing the composition of a committee in this system was remote. The random assignment checks carried out in Section 3 confirmed this assumption.

However, we have checked the robustness of our findings considering exclusively those competitions in which the evaluators were randomly selected from the whole body of full professors in the sub-field. In Table 9 we report the four main specifications reported in Table 3. Results are consistent with those found considering the all sample of competitions.

Table 9. Only sub-fields with completely randomly selected committees

	(1)	(2)	(3)	(4)
Female	-0.052***	-0.034***	-0.073***	-0.056***
	(0.018)	(0.017)	(0.023)	(0.022)
Female*(Females in Com.)			0.049	0.050
			(0.037)	(0.043)
Female*(% Females in Com.)				
Females in Com.			-0.012	-0.012
			(0.010)	(0.014)
% Females in Com.				
Relative Productivity		0.022***		0.021***
		(0.004)		(0.004)
Insider		0.237***		0.238***
		(0.035)		(0.035)
Connections		0.080***		0.080***
		(0.037)		(0.037)
University Job		0.004		0.004
		(0.024)		(0.024)
Age		-0.002		-0.002
		(0.002)		(0.002)
Observations	1387	1387	1387	1387
Pseudo R-squared	0.048	0.145	0.066	0.170

Notes: The Table reports marginal effects of Probit estimates (evaluated at the mean values of the explanatory variables in the sample). The dependent variable is *Success*. In all regressions we control for sub-field dummies, type of position dummy and the number of candidates. Standard errors (corrected for heteroskedasticity and clustered at the competition level) are reported in parentheses. The symbols ***, **, * indicate that coefficients are statistically significant, respectively, at the 1, 5, and 10 percent level.

9. Concluding Remarks

Females typically obtain worse results compared to their male counterparts in many dimensions of social and economic life. A large empirical evidence shows that females earn substantially less than men even

[7] For example, suppose that 10 vacancies are opened in a subfield, then 10*4*3=120 professors must be elected. Each committee is then composed selecting randomly four members from this pool of 120 elected members.

when they perform the same job and have the same qualification. In addition, the presence of females in top and high-ranking positions is negligible in many countries.

This state of affairs explains why policymakers and researchers often debate about what types of policies may promote gender equality. In the past, equal opportunities and equal treatment laws have been the main focus, reflecting the widespread idea that improvement in women's access to education would have allowed them to reach positions similar to those held by men. However, this kind of policy has not produced the expected results and unbalances in top and influential positions still persist. Recently, in order to overcome these inequalities many countries have introduced gender parity in top positions. However, whether the hiring or promotion of more women to influential positions represents an effective way to break the "glass ceiling" for females is still a matter of discussion (Chattopadhyay and Duflo 2004; Pande 2003; De Paola, Scoppa and Lombardo, 2010).

In this paper we have tried to shed some light on this issue focusing on female performance in academic promotions and trying to understand whether the gender of evaluators matters. Relying on a large randomized natural experiment consisting in the examinations for promotion to associate and full professor positions in the Italian University, where the allocation of evaluators to each competition was random, we have investigated the candidates' probability of success and how it is affected by the gender composition of evaluation committees.

From our analysis it emerges that, even after controlling for individual characteristics, measures of scientific productivity and indicators of social connections, females experiment a considerable lower probability of success. Interestingly, females' chances of success are affected by the gender of evaluators. In competitions in which the evaluators are exclusively males, female candidates suffer a reduction of their probability of success of about 6.4 percentage points: this implies that the probability of success of females is about 50% lower than males. On the other hand, gender discrimination almost vanishes when the candidates are judged by a mixed sex committee. We find very similar results across different type of positions and different fields.

Information on withdrawal decisions has allowed us to investigate whether our results are driven by self-fulfilling expectations or by different preferences for competition. Females may believe that they are less likely to succeed when the evaluation committee is composed exclusively by men and decide to retire from competition when facing an all-male committee. Nevertheless, we do not find evidence that the committee gender composition affects the candidates' probability of withdrawing from competition. In addition, the discrimination against females operated by all-male committees and the positive impact of mixed sex committees on female candidates' probability of success persists also when we exclude from our sample the candidates who have withdrawn from competition. This finding also suggests that gender differences in preferences for competition does not play an important role in explaining the worse performance of females.

Our findings showing that a greater number of females in influential positions may help other women to advance in their career are in line with those emerging from Zinovyeva and Bagues (2011)

showing, at least for competitions to full professor in Spain, a positive effect of female evaluators on the probability of success of female candidates, but are in contrast with the findings of Bagues and Esteve-Volart (2010) who, for the access to positions in the Spanish Judiciary, find that women in evaluation committees favor male candidates.

These diverging results suggest that the attitudes of mixed sex committees toward male and female candidates may change in relation to different contexts or in relation to the positions to be filled. Additional research seems necessary in order to better understand the role of females in preventing gender discrimination and to formulate policy recommendations aimed at promoting equality of treatment for men and women.

References

Albrecht, J., Bjorklund, A., and Vroman, S., (2003), "Is there a glass ceiling in Sweden?", *Journal of Labor Economics*, 21, pp. 145–177.

Altonji, J. and Blank, R., (1999), "Race and gender in the labor market", in Ashenfelter O. and Card, D., eds.), *Handbook of Labor Economics*. Volume 3C, pp. 3143–259, Amsterdam: North-Holland.

Arulampalam, W., Booth, A., and Bryan, L., (2007), "Is there a glass ceiling over Europe? Exploring the gender pay gap across the wages distribution", *Industrial and Labor Relations Review*, 60, pp. 163–186.

Bagues, M and Esteve-Volart, B., (2010), "Can Gender Parity Break the Glass Ceiling? Evidence from a Repeated Randomized Experiment", *Review of Economic Studies*, 77, pp. 1301-1328.

Bell, L., (2005), "Women-Led Firms and the Gender Gap in Top Executive Jobs", *IZA Discussion Paper* 1689.

Bertrand, M., (2011), "New Perspectives on Gender", *Handbook of Labor Economics*, vol. 4b.

Blackaby, D., Booth, A., and Frank, J., (2005), "Outside offers and the gender pay gap", *Economic Journal* 115, F81–F107.

Blau, F., and Kahn, M., (2003), "Understanding international differences in the gender pay gap", *Journal of Labor Economics*, 21, pp. 106–144.

Booth, A., (2009), "Gender and Competition", *Labour Economics*, 16, pp. 599–606.

Cahuc, P. and Zylberberg, A., (2004), *Labor Economics*, MIT Press.

Cameron, C., Gelbach, J. and Miller, D., (2006), "Robust Inference with Multi-way Clustering", *NBER Technical Working Papers* 0327.

Cardoso, A.R., and Winter-Ebmer, R., (2007), "Mentoring and segregation: female-led firms and gender wage policies", *IZA Discussion Paper* No. 3210, December.

Chattopadhyay, R. and Duflo E., (2004), "Women as Policy Makers: Evidence from a Randomized Policy Experiment in India", *Econometrica*, 72(5), 1409–1443.

Checchi, D., (1999), "Tenure. An Appraisal of a National Selection Process for Associate Professorship", *Giornale degli Economisti*, 58, 137-181.

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

De Paola, M., Scoppa, V., and Lombardo, R., (2010) "Can Gender Quotas Break Down Negative Stereotypes? Evidence from Changes in Electoral Rules", *Journal of Public Economics*, 94, pp. 344-353.

Ginther, D. and Hayes, K., (2003), "Gender Differences in Salary and Promotion for Faculty in the Humanities, 1977-1995", *Journal of Human Resources*, 38 (1), pp. 34-73.

Ginther, D. and Kahn, S., (2004), "Women in Economics: Moving Up or Falling Off the Academic Ladder", *Journal of Economic Perspectives*, Vol. 18(3), pp 193-214.

Ginther, D. and Kahn, S., (2009), "Does Science Promote Women? Evidence from Academia 1973-2001", NBER chapters in: *Science and Engineering Careers in the United States: An Analysis of Markets and Employment*, pp. 163-194.

Goldin, C., and Rouse, C., (2000), "Orchestrating impartiality: the impact of blind auditions on female musicians", *American Economic Review*, 90 (4), 715–741.

Hilmer, C., and Hilmer, M., (2010), "Are There Gender Differences in the Job Mobility Patterns of Academic Economists?", *American Economic Review*, 100, pp. 353-357.

Lavy, V., (2008), "Do Gender Stereotypes Reduce Girls' Human Capital Outcomes? Evidence from a Natural Experiment", *Journal of Public Economics*, Vol. 92(10-11), pp. 2083-2105.

Manning, A. and Saidi, F., (2010), "Understanding the Gender Pay Gap: What's Competition Got to Do with It?", *Industrial & Labor Relations Review*, 63(7).

McDowell, J. M., L., Singell, and Ziliak J. P., (1999), "Cracks in the Glass Ceiling: Gender and Promotion in the Economics Profession", *American Economic Review, Papers and Proceedings*, Vol. 89(2), pp. 392-396.

Pande, R., (2003), "Can Mandated Political Representation Increase Policy Influence for Disadvantaged Minorities?", *American Economic Review*, 93(4), 1132-1151.

Perotti, R., (2002), The Italian University System: Rules vs. Incentives. Presented at the first conference on "Monitoring Italy", ISAE, Rome.

Weichselbaumer, D., and Winter-Ebmer, R., (2005), "A Meta-Analysis of the International Gender Wage Gap", *Journal of Economic Surveys*, 19, pp. 479-511.

Wenneras, C. and Wold, A., (1997), "Nepotism and sexism in peer-review", *Nature*, Vol. 387, pp. 341-343.

Zinovyeva, N. and Bagues, M., (2011), "Does gender matter for academic promotion? Evidence from a randomized natural experiment", *IZA Discussion Paper* 5537.