Does gender matter for academic promotion: Evidence from a randomized natural experiment

Natalia Zinovyeva Manuel Bagues

IPP-CSIC Universidad Carlos III

Max Weber Multidisciplinary Research Workshop
February 2, 2011

Motivation

- In Europe in 2006 women accounted for 45% of PhD graduates, 36% of associate professors and a mere 18% of full professors. Similar pattern is observed in the US.

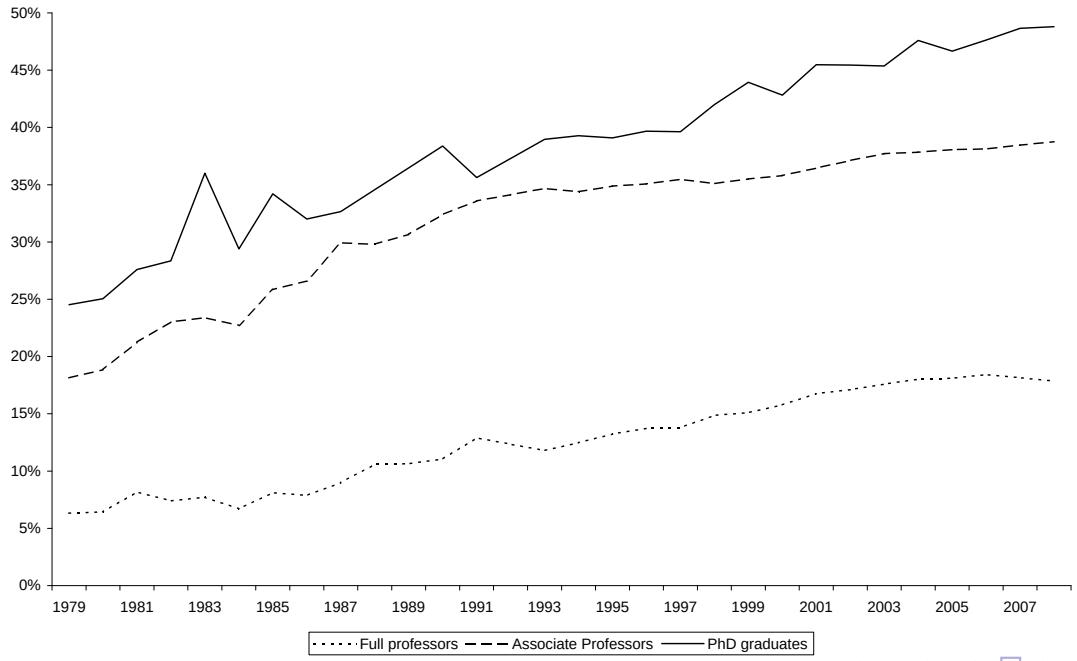
Motivation

- In Europe in 2006 women accounted for 45% of PhD graduates, 36% of associate professors and a mere 18% of full professors. Similar pattern is observed in the US.
- The pipeline seems to be ‘leaky’, particularly at the top.

Motivation

The presence of women among PhD graduates has grown from 36% to 49% over the last twenty years. In the same period, the incidence of women among faculty has increased from 30% to 39% among associate professors, but only from 11% to 18% among full professors.

Proportion of women in Spanish academia, by rank



Motivation

- In Europe in 2006 women accounted for 45% of PhD graduates, 36% of associate professors and a mere 18% of full professors. Similar pattern is observed in the US.
- The pipeline seems to be ‘leaky’, particularly at the top.
  - Family issues
  - Lack of role models in the upper echelons
  - Gender discrimination

Motivation

- Policy: gender quotas in scientific evaluation committees (Norway 1988, Finland 1995, Sweden 1999, Spain 2007)

Motivation

- Policy: gender quotas in scientific evaluation committees (Norway 1988, Finland 1995, Sweden 1999, Spain 2007)
- Empirical evidence on the effectiveness of gender quotas is scarce

Motivation

- Policy: gender quotas in scientific evaluation committees (Norway 1988, Finland 1995, Sweden 1999, Spain 2007)
- Empirical evidence on the effectiveness of gender quotas is scarce
  - Broder (1993) finds that female evaluators rate NSF female-authored grant proposals lower than do their male colleagues
  - Abrevaya and Hamermesh (2010) examine referee evaluations and do not find any effect of the interaction between the referees' and authors' gender
  - Bagues and Esteve-Volart (2010) find that female candidates to entry positions in the Spanish Judiciary are significantly less likely to be hired if they are (randomly) assigned to a committee with a relatively greater share of female evaluators
  - Booth and Leigh (2010) conduct an audit study in Australia and do not find any significant interaction between the gender of the applicant and the gender of the contact person in the hiring firm

Motivation

- Policy: gender quotas in scientific evaluation committees (Norway 1988, Finland 1995, Sweden 1999, Spain 2007)
- Empirical evidence on the effectiveness of gender quotas is scarce
  - Broder (1993) finds that female evaluators rate NSF female-authored grant proposals lower than do their male colleagues
  - Abrevaya and Hamermesh (2010) examine referee evaluations and do not find any effect of the interaction between the referees' and authors' gender
  - Bagues and Esteve-Volart (2010) find that female candidates to entry positions in the Spanish Judiciary are significantly less likely to be hired if they are (randomly) assigned to a committee with a relatively greater share of female evaluators
  - Booth and Leigh (2010) conduct an audit study in Australia and do not find any significant interaction between the gender of the applicant and the gender of the contact person in the hiring firm
- None of these studies considers promotion to top positions, where the phenomenon of glass ceiling is observed. More evidence is needed to understand the role of gender on the top of the career ladder.

Motivation

- We analyze how the gender composition of committees affects promotions to associate and full professor positions

Motivation

- We analyze how the gender composition of committees affects promotions to associate and full professor positions
- We consider Spanish national competitions (*habilitación*) that were held in the period 2002-2006

Motivation

- We analyze how the gender composition of committees affects promotions to associate and full professor positions
- We consider Spanish national competitions (*habilitación*) that were held in the period 2002-2006
- Main features of this paper:
  - Promotions to top-level positions
  - Large scale: all academic disciplines during five years
  - The identification strategy exploits the random assignment of evaluators to committees

Outline

Outline

National competitions (Habilitaciones)

- During 2002-2006 candidates to associate or full professor positions had to compete at the national level for a limited number of qualifications (similar system in France and Italy)
- Qualified candidates could apply for positions at the university level
- In practice, the subsequent competition was almost absent, because the number of qualifications was equal or lower than the number of all university openings

Procedure

- The Ministry published the call for exams

Procedure

- The Ministry published the call for exams
- In the following 20 days candidates could apply

Procedure

- The Ministry published the call for exams
- In the following 20 days candidates could apply
- Once the list of eligible candidates was formed, evaluators (seven per committee) were selected by random draw out of the list of eligible evaluators:
  - In full professor exams all evaluators were full professors
  - In associate professor exams committees were composed of three full professors and four associate professors

Content of exams

- Exams to associate professor positions:
  1. CV presentation
  2. Lecture on one (randomly chosen) topic out of the syllabus
  3. Research seminar
- Exams to full professor positions:
  1. CV presentation
  2. Research seminar

Outline

The sample

- Overall 1014 examinations in 188 disciplines were held during 2002-2006
- We exclude exams in which the number of potential evaluators was not big enough to form a committee and in which there were no females among potential evaluators
- The final sample includes 455 exams to associate professor and 436 exams to full professor

Individual information

- Eligible evaluators: 21,945 associate professors and 7,909 full professors
- Applicants: 18,792 applications (9,952 candidates) to associate professor and 13,224 applications (6,037 candidates) to full professor
- For all eligible evaluators and candidates:
  - Gender
  - Age
  - ISI publications and received citation (except Humanities)
- For all eligible evaluators and candidates to full professor positions:
  - Affiliation
  - Tenure

Outline

Identification

- We compare the outcomes of those exams, in which the expected share of females in the evaluation committee μ_e was the same, but the realizations of random draws resulted in committees with different gender composition s_e: $E[s_e] - \mu_e = E[s_e|\cdot] - \mu_e = 0$.

Expected and actual committee composition

	1	2	3	4
	Associate professor exams		**Full professor exams**	
	Expected Committees	Actual Committees	Expected Committees	Actual Committees
Share female	0.26	0.26	0.15	0.15
	[0.01]	[0.01]	[0.01]	[0.01]
Zero female	0.18	0.17	0.36	0.36
	[0.01]	[0.02]	[0.01]	[0.02]
One female	0.28	0.28	0.36	0.35
	[0.01]	[0.02]	[0.01]	[0.02]
Two female	0.25	0.27	0.19	0.18
	[0.01]	[0.02]	[0.01]	[0.02]
Three female	0.17	0.18	0.07	0.08
	[0.01]	[0.02]	[0.01]	[0.01]
Four or more female	0.12	0.10	0.03	0.03
	[0.01]	[0.01]	[0.00]	[0.01]
Age	47.56	48.03	51.14	51.19
	[0.13]	[0.19]	[0.10]	[0.19]
Publications, weighted by co-authors	6.70	6.61	8.96	8.45
	[0.24]	[0.27]	[0.36]	[0.38]
Citations per publication	11.13	11.12	12.26	11.66
	[0.25]	[0.32]	[0.28]	[0.32]

Notes: The sample includes 455 exams to associate professor positions and 436 exams to full professor positions. Standard errors are reported in square brackets.

Estimation model

- Baseline model:

$$y_{ie} = \beta_0 + \beta_1 f_i + \beta_2 (s_e - \mu_e) + \beta_3 f_i (s_e - \mu_e) + \beta_4 z_e + \epsilon_{ie} \quad (1)$$

- where:
  - y_{ie} indicates whether individual i qualified in exam e
  - f_i is an indicator for females
  - $s_e - \mu_e$ is the difference between the actual and the expected share of female evaluators in committee e
  - z_e is the number of available positions per candidate

The effect of committees' gender composition on candidates' success - Full professor exams

	1	2	3	4	5	6	7
	Full professor exams						
	All	Engineering	Math and Physics	Medicine	Social Sciences	Biology and Chemistry	Humanities
Female candidate	-0.013**	0.005	-0.012	-0.022	-0.020	-0.014	-0.013
	(0.005)	(0.020)	(0.013)	(0.019)	(0.016)	(0.014)	(0.008)
Proportion of female evaluators	-0.042***	-0.076	0.002	-0.119**	-0.061	0.002	-0.050*
	(0.016)	(0.063)	(0.030)	(0.047)	(0.048)	(0.040)	(0.026)
Female candidate*Proportion of female evaluators	0.127***	0.449	-0.009	0.347***	0.187	0.017	0.123*
	(0.046)	(0.342)	(0.139)	(0.127)	(0.123)	(0.111)	(0.071)
Positions per candidate	0.965***	1.017***	0.999***	1.001***	0.984***	0.956***	0.950***
	(0.003)	(0.085)	(0.002)	(0.003)	(0.014)	(0.011)	(0.016)
Constant	0.102***	0.091***	0.083***	0.101***	0.115***	0.113***	0.102***
	(0.001)	(0.004)	(0.002)	(0.005)	(0.005)	(0.004)	(0.003)
Adjusted R-squared	0.028	0.014	0.020	0.033	0.024	0.028	0.035
Number of observations	13224	1214	1877	1473	1603	2511	4546

Notes: OLS estimates. Standard errors are clustered by applicant and by exam. '(Proportion of female evaluators)' is the difference between the actual and the expected proportion of female evaluators in the committee.

The effect of committees' gender composition on candidates' success - Associate professor exams

	1	2	3	4	5	6	7
	Associate professor exams						
	All	Engineering	Math and Physics	Medicine	Social Sciences	Biology and Chemistry	Humanities
Female candidate	-0.020***	-0.003	-0.009	-0.008	-0.025*	-0.058***	-0.023***
	(0.004)	(0.011)	(0.009)	(0.012)	(0.015)	(0.012)	(0.008)
Proportion of female evaluators	0.015	0.004	0.011	0.036	0.041	-0.013	0.035
	(0.012)	(0.024)	(0.009)	(0.049)	(0.042)	(0.044)	(0.028)
Female candidate*Proportion of female evaluators	-0.050*	-0.075	-0.053	-0.048	-0.090	-0.009	-0.072
	(0.026)	(0.096)	(0.036)	(0.108)	(0.082)	(0.089)	(0.048)
Positions per candidate	0.926***	0.943***	0.994***	1.014***	0.828***	0.879***	0.927***
	(0.022)	(0.026)	(0.018)	(0.042)	(0.143)	(0.049)	(0.022)
Constant	0.119***	0.113***	0.082***	0.098***	0.118***	0.168***	0.141***
	(0.002)	(0.002)	(0.003)	(0.004)	(0.008)	(0.007)	(0.004)
Adjusted R-squared	0.033	0.030	0.030	0.034	0.024	0.035	0.030
Number of observations	18792	3028	3260	2342	2689	1929	5544

Notes: OLS estimates. Standard errors are clustered by applicant and by exam. '(Proportion of female evaluators)' is the difference between the actual and the expected proportion of female evaluators in the committee.

The rank of evaluators

The difference in results across the two types of exams might potentially reflect differences in the behavior of evaluators associated with their rank (recall the differences in the design of committees). However, we reject this possibility:

	Associate professor exams
Female candidate	-0.020***
	(0.004)
Proportion of female full-professor evaluators	0.003
	(0.026)
Female candidate*Proportion of female full-professor evaluators	-0.023
	(0.058)
Proportion of female associate-professor evaluators	0.020
	(0.015)
Female candidate*Proportion of female associate-professor evaluators	-0.061*
	(0.033)
Constant	0.120***
	(0.002)
Adjusted R-squared	0.033
Number of observations	18792

Other committee characteristics

- We test that above results cannot be attributed to other committee characteristics
  - Age
  - Tenure
  - Research quality

Non-linearities

- What if the effect of the gender committee composition is not linear?
  - Maybe the presence of at least one female evaluator in the committee affects the voting behavior of male evaluators
  - Given that decisions are taken on a majority basis, the presence of female majority might be particularly strong
- We estimate non-linear effects of committee composition:

$$y_{ie} = \gamma_0 + \sum_{j=1}^{7} \gamma_j d_{je} + \lambda_0 f_i + \sum_{j=1}^{7} \lambda_j f_i d_{je}$$
$$+ \sum_{j=1}^{7} \delta_j p(d_{je}) + \sum_{j=1}^{7} \mu_j f_i p(d_{je}) + \nu z_e + \epsilon_{ie} \quad (2)$$

where d_{je} is a dummy that takes value one if the number of females in committee e is equal to j and $p(d_{je})$ is the probability that exactly j female evaluators were drawn.

$\Rightarrow$ The data cannot reject hypothesis of the linearity of the effect

Which committees discriminate?

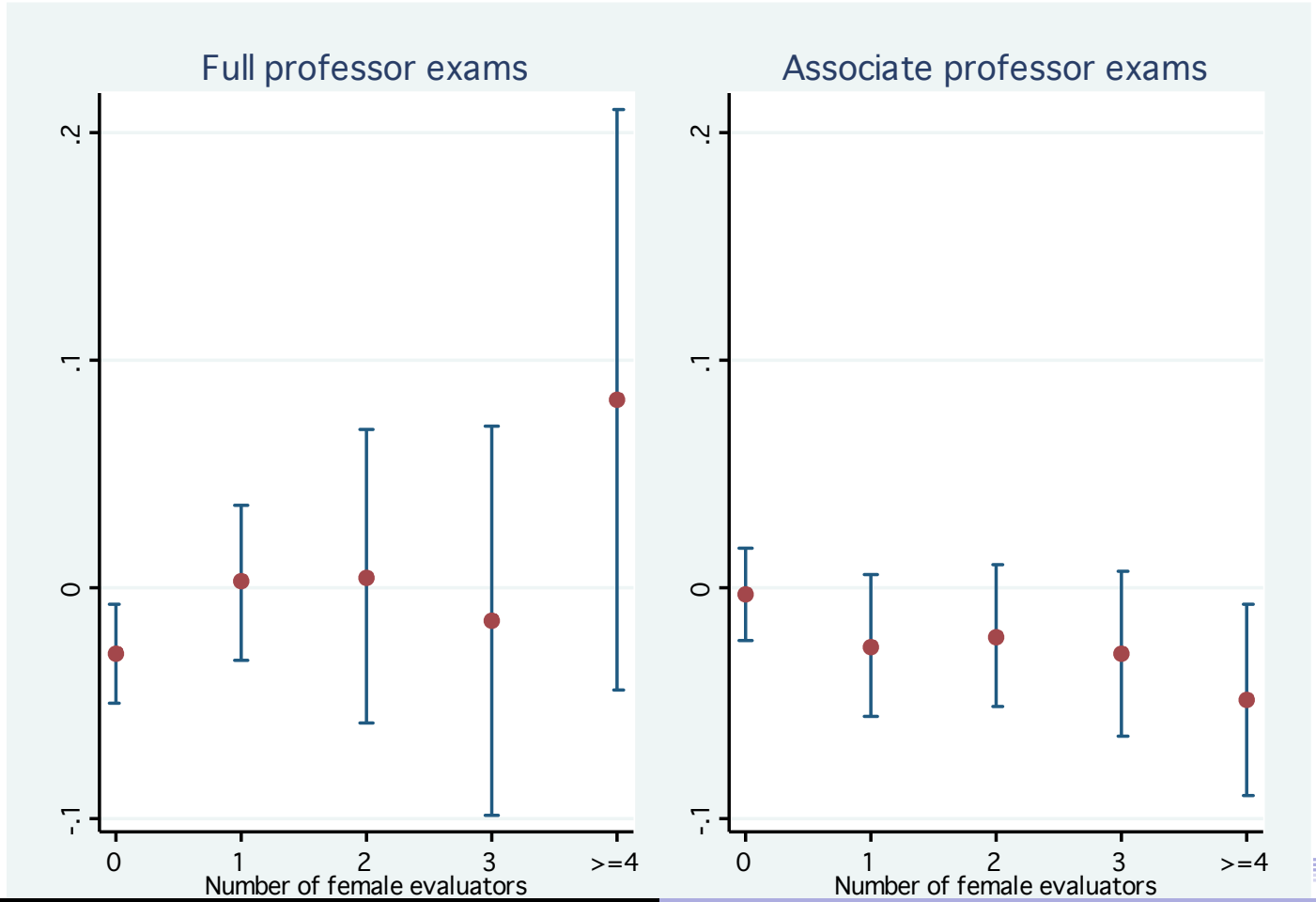
- Baseline model + individual quality controls

$$y_{ie} = \gamma_0' + \sum_{j=1}^{7} \gamma_j' d_{je} + \lambda_0' f_i + \sum_{j=1}^{7} \lambda_j' f_i d_{je}$$
$$+ \sum_{j=1}^{7} \delta_j' p(d_{je}) + \sum_{j=1}^{7} \mu_j' f_i p(d_{je}) + \nu' z_e + \eta q_{ie} + \epsilon_{ie} \quad (3)$$

  where vector $q_i e$ proxies the quality of candidate i sitting exam e.
- We use information on research output: (i) the number of ISI publications weighted by the number of co-authors, and (ii) received citations per publication. We also control for candidates' age. All measures are normalized for candidates in the same exam.
- As long as q_{ie} captures all gender differences in candidates' quality, $\lambda_j', j \in \{0, \ldots, 7\}$, provide information on whether female candidates are being discriminated against by committees with j female evaluators. .

Which committees discriminate?

The point and interval estimates of gender differences among comparable applicants, by committee composition:



Outline

Differences in evaluation or differences in performance?

1. Some candidates may decide not to take the exam if, given the observed committee composition, their expected probability of being promoted is not high enough to compensate the cost of attending the exam.
   - Evidence is not consistent with this hypothesis: there are no differences in the effect of committee gender composition on the success rate across candidates with, respectively, lower and higher cost of attending the exam as measured by travel time.

Differences in evaluation or differences in performance?

1. Some candidates may decide not to take the exam if, given the observed committee composition, their expected probability of being promoted is not high enough to compensate the cost of attending the exam.
   - Evidence is not consistent with this hypothesis: there are no differences in the effect of committee gender composition on the success rate across candidates with, respectively, lower and higher cost of attending the exam as measured by travel time.
2. The composition of the committee could affect the performance of some candidates during the exam ("stereotype threat").
   - The possibility of the "stereotype threat" playing a role might have been larger in exams to associate professor positions.
   - However, the evidence shows that female applicants to associate professor positions are relatively more successful, not less, in committees with more men.

Why are male evaluators relatively less favorable towards female candidates to full professor?

1. Information asymmetries across genders
   - Information asymmetries across genders might arise if genders specialize in different subfields or if they interact through gendered networks.
   - Testable implication: Information asymmetries should be less severe in smaller disciplines.
   - Rejected by the data

Why are male evaluators relatively less favorable towards female candidates to full professor?

1. Information asymmetries across genders
   - Information asymmetries across genders might arise if genders specialize in different subfields or if they interact through gendered networks.
   - Testable implication: Information asymmetries should be less severe in smaller disciplines.
   - Rejected by the data
2. Age-based discrimination: male evaluators may be biased against old females but not against young females.
   - Testable implication: Older female candidates should be relatively less likely (than younger female candidates) to be promoted in committees with more male evaluators.
   - Not supported by the data

Why are male evaluators relatively less favorable towards female candidates to full professor?

1. Information asymmetries across genders
   - Information asymmetries across genders might arise if genders specialize in different subfields or if they interact through gendered networks.
   - Testable implication: Information asymmetries should be less severe in smaller disciplines.
   - Rejected by the data
2. Age-based discrimination: male evaluators may be biased against old females but not against young females.
   - Testable implication: Older female candidates should be relatively less likely (than younger female candidates) to be promoted in committees with more male evaluators.
   - Not supported by the data
3. Cronyism: evaluators tend to favor their friends, and friends tend to be of the same gender.
   - Testable implication: The effect of the gender committee composition in full professor exams should be at least partially attributed to the gendered pattern of academic inbreeding.
   - Not supported by the data

Why are male evaluators relatively less favorable towards female candidates to full professor?

1. Information asymmetries across genders
   - Information asymmetries across genders might arise if genders specialize in different subfields or if they interact through gendered networks.
   - Testable implication: Information asymmetries should be less severe in smaller disciplines.
   - Rejected by the data
2. Age-based discrimination: male evaluators may be biased against old females but not against young females.
   - Testable implication: Older female candidates should be relatively less likely (than younger female candidates) to be promoted in committees with more male evaluators.
   - Not supported by the data
3. Cronyism: evaluators tend to favor their friends, and friends tend to be of the same gender.
   - Testable implication: The effect of the gender committee composition in full professor exams should be at least partially attributed to the gendered pattern of academic inbreeding.
   - Not supported by the data
4. Ambivalent sexism (Glick and Fiske, 1996): Male evaluators may dislike having females on top positions, but not at lower level positions.

Outline

Conclusions

- We find that gender committee composition affects the gender of promoted candidates:
  - In full professor exams, an additional male evaluator reduces by 14% the number of females that qualify.
  - When these committee members decide on promotions to associate professor positions, no significant interaction between evaluators' and candidates' gender could be observed.
- Explanation?
  - Male evaluators exhibit ambivalent sexism: they discriminate female candidates only when they apply to top positions.
- Policy implication:
  - The introduction of gender quotas may have a positive effect on the number of promoted female candidates *only* at the upper levels of the career ladder. The presence of a single female on the committee is enough to overcome the gender bias in evaluation.
  - Existent gender quotas should be reconsidered: given that sitting on committees reduces the available time for research, they might lower the productivity of women who have managed to overcome the glass ceiling.

Conclusions

- Companion paper: Zinovyeva (2010) studies the optimal composition of committees according to other dimensions. The paper finds that the best candidates tend to be promoted when:

Conclusions

- Companion paper: Zinovyeva (2010) studies the optimal composition of committees according to other dimensions. The paper finds that the best candidates tend to be promoted when:
  - Evaluators are good researchers

Conclusions

- Companion paper: Zinovyeva (2010) studies the optimal composition of committees according to other dimensions. The paper finds that the best candidates tend to be promoted when:
  - Evaluators are good researchers
  - Evaluators do not belong to the same institution as candidates