Does the Gender Composition of Scientific Committees Matter?

Manuel Bagues
Aalto University

Mauro Sylos Labini
University of Pisa

Natalia Zinovyeva
Aalto University

Few women in top positions in academia

- In Europe, 47% of PhD graduates, 37% of associate professors and 21% of full professors are women (European Commission 2016). Similar patterns in the US and Japan.
  - Germany: 45% of PhD graduates and 17% of FPs

Few women in top positions in academia

- Pipeline theory
  - CSWEP Newsletter (2012), Bagues (2013)
- Differences in productivity
  - Lack of mentors, role models and/or research networks: CeMENT (Blau et al. 2010)
  - Family: Ginther and Kahn (2004)
  - Non-promotable tasks: Vesterlund, Babcock and Weingart (2014)
- Behavior in the labor market
  - Women do not apply for promotions: Bosquet, Combes and Garcia-Penalosa (2013), De Paola, Ponzo and Scoppa (2014)
  - Women shy away from competitions: Niederle and Vesterlund (2007), Buser, Niederle and Oosterbeek (2013)
  - Poor bargaining: Babcock and Laschever (2009); Blackaby, Booth and Frank (2005)

Demand-side explanation: discrimination by (mostly male) evaluators

1. Gender segregation across fields combined with same field preference
   - Dolado et al. 2012, Hale and Regev 2011
2. Old boys networks
   - Zinovyeva and Bagues 2015, Bagues, Sylos-Labini and Zinovyeva 2014, Combes, Linnemer and Visser 2008
3. Negative stereotypes held by men (World Value Survey)
4. Lack of diversity in committees affecting negatively the quality of evaluations (Woolley et al. 2010)

- Gender quotas in scientific committees:
  - Finland (1995), Spain (2007), France (2014)
  - European Commission
  - Internal guidelines in many universities

Are quotas effective?

- Quotas are costly for senior female researchers
- More women in committees may may not benefit female candidates
  - Female evaluators may be similar to male (Mendez and Busenbark 2012)
  - Female evaluators may be not influential in committees (Karpowitz et al. 2012, Brescoll 2011)
  - Men may feel licensed to discriminate and/or simply express more honest opinions about female candidates (Khan and Dhar 2006, Monin and Miller 2011)
  - Backlash effect: men do not always respond favorably to the presence of gender diversity in domains that are historically male-dominated (Crocker and McGraw 1984)

Empirical evidence

- Existent evidence: mixed results
  - Same-sex preference
    - Casadevall and Handelsman (2013, IBO), De Paola and Scoppa (2014, IBO)
  - Opposite-sex preference:
    - Broder (1993, FE), Ellemers et al. (2004, IBO)
  - Gender of evaluators has no statistically significant effect:
    - Moss-Racusin et al. (2012, RCT), Steinpreis et al. (1999, RCT), Abrevaya and Hamermesh (2012, FE), Jayasinghe et al. (2003, FE), Milkman, Akinola and Chugh (2015, RCT), Williams and Ceci (2015, RCT)
- No empirical evidence on specific mechanisms suggested by the theory

Evidence from two large-scale randomized experiments

Nation-wide qualification exams in Italy and Spain

- Promotions to an associate or full professorship require a qualification granted by a centralized committee
- Large-scale: two countries, several years, every field, two different positions
  - 100,000 applications, 8,000 evaluators
- Randomized natural experiments: Evaluators selected using a random draw
- Detailed information on evaluators' and candidates'
  - Research production
  - Academic connections
  - Field of specialization
- Information about individual voting behavior within committees (Italy)
- Information about future performance of candidates (Spain)

Preview of main findings

- Do more women in committee increase the chances of female candidates?
  - No!
- Individual voting within committees:
  - Female evaluators are slightly more favorable towards female candidates (than male evaluators)
  - The presence of female evaluators makes male evaluators tougher upon female candidates.
- Does the gender composition of committees affect the quality of promoted candidates?
  - No!
- Why don't we observe a stronger same-sex preference?
  - Old-boys networks
  - Gender segregation across research interests
- Stereotypes
  - Gender only matters when evaluators are not familiar with candidates' research

Outline

Outline

Institutional Background

- Nation-wide evaluations to become associate or full professor (1^{st} stage):
  - In Italy, *Abilitazione Scientifica Nazionale* (2012-2014)
  - In Spain, *Habilitación* (2002-2006)

- The timeline of the national evaluations:
  1. The call is announced
  2. Candidates apply
  3. Random selection of evaluators that satisfy minimum requirements
  4. Evaluation takes place

Italy vs. Spain

- In Italy:
  - Evaluations on CVs and 'nudge' on bibliometric indicators
    - *Sciences*: Publications, citations, h-index
    - *SS&H*: Top publications, all publications, books
  - No limit on the number of qualifications
  - Committee members: 5 FP in all exams, 4 Italians + 1 Foreign
  - Promotion requires 4 votes (out of 5)
  - Very transparent: CVs, evaluation criteria and evaluations published on-line
- In Spain:
  - Oral qualifying exams: 2 qualifying stages in FP exams, 3 in AP exams
  - Number of qualifications limited (tournament)
  - Committee members: 7 FP in FP exams, 3 FP + 4 AP in AP exams
  - Promotion requires 4 votes (out of 7)
  - Only final outcome observed

Outline

Evaluators

- In Italy:
  - 184 committees in corresponding fields
  - 7,241 eligible FP (80% based in Italy, 20% based abroad)
  - Women in committees 19%
    - all-male committees 41%; female majority 8%.
    - women younger, shorter CV
  - 8% of initially rostered evaluators resigned
  - gender differences
- In Spain:
  - 967 committees in 174 fields.
  - 7,955 eligible FP and 21,975 eligible AP
  - Women in committees 19%
    - all-male committees 31%; female majority 6%.
    - women younger, shorter CV
  - 2% of initially rostered evaluators resigned

Candidates

- In Italy, 69,020 initial applications, 375 per committee, 38% women
  - Quality measures: publications, books, chapters, conference proceedings, patents, Article Influence Score, length of CV.
  - Other individual characteristics: age, gender, type of contract, affiliation, field of research, application order
  - 14% of candidates dropped out; 59,150 final candidates
- In Spain, 31,243 applications, 32 candidates per exam, 34% women
  - Quality measures: International articles (ISI), Spanish articles, books and chapters (Dialnet), patents (EPO), PhD students advised and doctoral theses committee participation (TESEO)
  - Other characteristics: age, gender, affiliation, field of interest (TESEO)

Table: **Descriptive statistics**

	1	2	3	4	5	6
	Spain			**Italy**		
	Male	Female	p-value	Male	Female	p-value
Age	0.01	-0.02	0.005	0.01	-0.02	0.000
All Publications:	0.05	-0.10	0.000	0.04	-0.07	0.000
- Articles	0.05	-0.10	0.000	0.07	-0.11	0.000
- Books	0.01	-0.02	0.000	0.05	-0.08	0.000
- Book chapters	0.01	-0.01	0.005	0.01	-0.02	0.000
- Patents	0.00	-0.01	0.060	0.02	-0.04	0.000
Average AIS (Sciences)	0.01	-0.03	0.022	0.03	-0.06	0.000
A-journal articles (SSH)	0.05	-0.07	0.000	0.04	-0.05	0.000
PhD students advised	0.03	-0.06	0.000			
PhD committees	0.04	-0.08	0.000			
Qualified	0.12	0.11	0.002	0.38	0.35	0.000

Connections and research similarity

- Strong ties
  - Colleagues (2.8% in Italy and 4.6% in Spain)
  - Coauthors (1.4% in Italy and 0.4% in Spain)
  - Student-advisor relationship (0.2% in Spain)
- Weak ties
  - Participation in assessment of the same doctoral thesis (1.3% in Spain)
- Research interest overlap
  - Same officially defined subfield for tenured researchers (60% in Italy)
  - Overlap of Unesco subfield codes of doctoral dissertations (Spain):

$$Overlap_{ij} = \frac{S_i S'_j}{(S_i S'_i)^{1/2}(S_j S'_j)^{1/2}}$$

where $S_i = (S_{1i}...S_{Ci})$ and $S_j = (S_{1j}...S_{Cj})$, and S_{Ci} is the share of dissertations in category C in which individual i has been involved.

Outline

Causal effect of committee gender composition

Intention-to-treat effect

- We estimate the following equation using the sample of initial applicants and evaluators:

$$Y_{i,e} = \beta_1 Female_i + \beta_2 Female_e^{initial} + \beta_3 Female_i * Female_e^{initial} + \\ + \beta_4 Female_e^{expected} + \beta_5 Female_i * Female_e^{expected} + \mathbf{X_i}\beta_\mathbf{6} + \epsilon_{i,e} \quad (1)$$

  where
  - $Female_i$ is an indicator for female candidates
  - $Female_e^{initial}$ is the proportion of women in the committee initially drawn
  - $Female_e^{expected}$ is the expected proportion of women in the committee
  - $\mathbf{X_i}$ individual predetermined characteristics
  - Standard errors are clustered at the committee level
- β_3 captures the causal impact of committees' initial gender composition on the relative success rate of female candidates
  - Key identification assumption: random selection of committee members

Causal effect of committee gender composition

Instrumental variables estimation

- 2nd stage:

$$Y_{i,e} = \beta_1 Female_i + \beta_2 \widehat{Female}_e^{final} + \beta_3 Female_i * \widehat{Female}_e^{final} + \\ + \beta_4 Female_e^{expected} + \beta_5 Female_i * Female_e^{expected} + \mathbf{X_i}\beta_{\mathbf{6}} + \epsilon_{i,e} \quad (2)$$

- 1st stage:

$$Female_e^{final} = a_1 Female_i + a_2 Female_e^{initial} + a_3 Female_i * Female_e^{initial} + \\ + a_4 Female_e^{expected} + a_5 Female_i * Female_e^{expected} + \mathbf{X_i a_6} + u_{i,e} \quad (3)$$

$$Female_i * Female_e^{final} = b_1 Female_i + b_2 Female_e^{initial} + b_3 Female_i * Female_e^{ini} \\ + b_4 Female_e^{expected} + b_5 Female_i * Female_e^{expected} + \\ + \mathbf{X_i b_6} + v_{i,e} \quad (4)$$

Table: **Randomization check**

	1	2	3	4	5	6	7	8	9	10	11
Dependent variable:	All Publications	Articles	Books	Chapters	Patents	Total AIS	A-journal articles	Coauthors per article	Proportion first-author	Proportion last-author	Age
						Italy					
Share of women in committee	0.014	0.005	-0.023	-0.020	0.019	-0.005	0.038	0.017	-0.040	-0.011	-0.065
	(0.034)	(0.031)	(0.031)	(0.028)	(0.021)	(0.030)	(0.027)	(0.031)	(0.023)	(0.041)	(0.035)
Female candidate*Share of women in committee	-0.027	-0.001	0.059	0.048	-0.040	0.018	-0.087	-0.044	0.093	0.031	0.150
	(0.079)	(0.071)	(0.066)	(0.063)	(0.049)	(0.064)	(0.061)	(0.067)	(0.053)	(0.088)	(0.080)
						Spain					
Share of women in committee	-0.019	-0.029	0.004	0.002	0.024	-0.068	-0.023	-0.040	-0.020	-0.023	0.034
	(0.030)	(0.031)	(0.022)	(0.022)	(0.015)	(0.027)	(0.022)	(0.030)	(0.031)	(0.032)	(0.034)
Female candidate*Share of women in committee	0.015	0.038	-0.017	-0.010	-0.062	0.152	0.043	0.103	0.045	0.042	-0.093
	(0.077)	(0.078)	(0.054)	(0.055)	(0.038)	(0.068)	(0.057)	(0.076)	(0.078)	(0.080)	(0.086)

Dependent variable:	Qualified						Applied
	OLS	OLS	ITT	IV	IV	IV	IV
				Italy			
Female candidate	-0.028	-0.015	-0.004	0.001	0.008	0.009	-0.026
	(0.006)	(0.005)	(0.009)	(0.011)	(0.008)	(0.007)	(0.006)
Share of women in committee			0.000	-0.0004	-	-	-
			(0.059)	(0.071)			
Female candidate* Share of women in committee			-0.092	-0.116	-0.128	-0.132	-0.025
			(0.036)	(0.050)	(0.035)	(0.036)	(0.026)
Number of observations	69020	69020	69020	69020	69020	69020	69020
				Spain			
Female candidate	-0.022	-0.014	-0.009	-0.009	-0.011	-0.010	
	(0.004)	(0.004)	(0.007)	(0.007)	(0.007)	(0.007)	
Share of women in committee			0.011	0.012	-	-	
			(0.017)	(0.018)			
Female candidate* Share of women in committee			-0.018	-0.019	-0.016	-0.022	
			(0.026)	(0.027)	(0.028)	(0.028)	
Number of observations	31243	31243	31243	31243	31243	31243	
Controls for both panels:							
Candidate characteristics		Yes	Yes	Yes	Yes	Yes	Yes
Exam FE	Yes	Yes			Yes	Yes	Yes
Expected share of women			Yes	Yes	Yes	Yes	Yes
Female candidate* Exp. share women			Yes	Yes	Yes	Yes	Yes
Committee characteristics						Yes	

Do female evaluators increase the *quality of selection*?

- Compare the observable quality of candidates who qualified in committees with different gender compositions:

$$x_{ie} = \beta_0 + \beta_1 Female_e^{final} + \beta_2 Female_e^{expected} + \epsilon_{ie}$$

where x_{ie} proxy of candidate i's quality, measured at the time of the evaluation or during the following five years

$Female_e^{expected}$ is a set of non-parametric controls for the *expected share of women* in the committee.

Table: **Quality of qualified candidates**

	1	2	3	4	5	6	7
Dep. var.:	Publications	Citations	Total AIS	A-journal articles	PhD students advised	PhD thesis committees	Success in future evaluations
	A. Italy, before the evaluation						
All	0.017	0.130	-0.055	-0.135			
	(0.088)	(0.117)	(0.157)	(0.255)			
Women	-0.044	0.139	0.154	-0.102			
	(0.112)	(0.143)	(0.170)	(0.317)			
Men	0.029	0.098	-0.208	-0.213			
	(0.101)	(0.150)	(0.211)	(0.251)			
	B. Spain, before the evaluation						
All	0.022	0.072	-0.088	-0.200	0.125	-0.147	
	(0.145)	(0.223)	(0.244)	(0.237)	(0.136)	(0.132)	
Women	0.210	0.469	-0.004	-0.142	0.580	0.053	
	(0.206)	(0.370)	(0.399)	(0.329)	(0.229)	(0.220)	
Men	-0.124	-0.242	-0.215	-0.219	-0.170	-0.303	
	(0.193)	(0.291)	(0.301)	(0.333)	(0.176)	(0.168)	
	C. Spain, after the evaluation						
All	0.016	-0.060	-0.098	-0.173	0.175	-0.086	0.042
	(0.132)	(0.218)	(0.227)	(0.181)	(0.135)	(0.136)	(0.052)
Women	0.345	-0.009	-0.102	0.170	0.119	-0.117	0.001
	(0.213)	(0.356)	(0.376)	(0.288)	(0.212)	(0.231)	(0.054)
Men	-0.187	-0.140	-0.247	-0.266	0.080	-0.134	0.019
	(0.182)	(0.281)	(0.284)	(0.252)	(0.191)	(0.186)	(0.077)

Information from individual votes

- Female candidates are relatively less successful when being evaluated by mixed-gender committees:
  1. Women vote against women?
  2. Or, in the presence of women, men are less favorable towards women?

Women vote against women?

- First, we compare the assessments of male and female evaluators sitting in the same committee. We estimate the following equation:

$$V_{ije} = \beta_0 + \beta_1 Female_j + \beta_2 Female_i * Female_j + \mu_{ie} + \epsilon_{ije}$$

where:

  - V_{ije}: takes value one if evaluator j casted a positive vote for candidate i in evaluation e.
  - $Female_j$: indicator for female evaluators
  - $Female_i$: indicator for female candidates
  - μ_{ie}: set of application fixed effects
  - Standard errors are clustered at the committee level

Table: **Individual voting**

Female evaluator	-0.001
	(0.007)
Female evaluator * Female candidate	0.007
	(0.005)
Application FE	Yes
Adj. R-squared	0.846
N	294,656

Does the presence of women in the committee affect male evaluators' voting behavior?

- We compare the voting behavior of male evaluators in all-male committees and in mixed-gender committees

$$V_{ije} = \beta_0 + \beta_1 Female_i + \beta_2 Female_{je}^{final} + \beta_3 Female_i * Female_{je}^{final} + \\ + \beta_4 Female_{je}^{expected} + \beta_5 Female_i * Female_{je}^{expected} + \mathbf{X_i}\beta_4 + \epsilon_{ij}$$

where:

- V_{ije}: takes value one if evaluator j casted a positive vote for candidate i in evaluation e.
- $Female_i$: indicator for female candidate
- $Female_{je}$: share of women in the committee of evaluator j
- $Female_{je}^{expected}$: expected share of women in the committee of evaluator j
- Standard errors are clustered at the committee level

Does the presence of women in the committee affect male evaluators' voting behavior?

- We compare the voting behavior of male evaluators in all-male committees and in mixed-gender committees

$$V_{ije} = \beta_0 + \beta_1 Female_i + \beta_2 Female_{je}^{final} + \beta_3 Female_i * Female_{je}^{final} + \\ + \beta_4 Female_{je}^{expected} + \beta_5 Female_i * Female_{je}^{expected} + \mathbf{X_i}\beta_4 + \epsilon_{ij}$$

- Identification assumptions:
  1. Random selection of committee members
  2. Evaluators' attrition is as good as random
     - 8% of initially rostered evaluators resigned
  3. Candidates' withdrawal is as good as random
     - 15% initial applicants withdrew their application

Table: **Individual voting**

	1	2	3
	Male evaluators		
Female candidate	-0.0004	0.008	-0.003
	(0.009)	(0.006)	(0.006)
Share of women in committee	0.017	-	-
	(0.079)		
Female candidate*Share of women in committee	-0.042	-0.061	-0.078
	(0.043)	(0.030)	(0.030)
Controls:			
Application FE	Yes		
Expected share of women	Yes	Yes	Yes
Female candidate*Expected share of women	Yes	Yes	Yes
Candidate characteristics	Yes	Yes	Yes
Exam FE		Yes	Yes
Number of observations	240988	240988	281289

Outline

Mechanisms

1. Why does the presence of women in the committee affect the voting behavior of male evaluators?
2. Why are women not more supportive of other women?

Mechanisms

- Why does the presence of women in the committee affect the voting behavior of male evaluators?
  1. backlash against female candidates

Table: **Degree of feminization**

	1	2	3	4
	Italy		Spain	
Feminization of field	$> median$	$< median$	$> median$	$< median$
	-0.149***	-0.072	-0.018	-0.016
	(0.042)	(0.057)	(0.040)	(0.037)

Notes: IV estimates. Standard errors are clustered by exam.

Mechanisms

- Why does the presence of women in the committee affect the voting behavior of male evaluators?
  1. backlash against female candidates
  2. licensing effect
  3. male identity priming

Mechanisms

- Women favor women, but the effect is not economically or statistically significant. Why?
  1. 'Old boys' network
  2. Gender segregation across subfields
  3. Stereotypes

'Old boys' network

- 'Old boys' network $\rightarrow$ same-sex preference
  - Gendered networks ✓
  - Connection premium ✓
  - Connections in committee ×

'Old boys' network

- Are networks gendered?

$$Link_{ij} = \beta_0 + \beta_1 Female_i + \beta_2 Female_j + \\ + \beta_3 Female_i * Female_j + \mathbf{D_c}\beta_\mathbf{4} + \epsilon_{ij},$$

where:

- unit of observation: candidate x eligible evaluator in the field
- $Female_i$: indicator for female candidate
- $Female_j$: indication for female eligible evaluator
- D_c: field fixed-effects

Table: **Gender pattern of links between candidates and eligible evaluators**

	1	2	3	4	5	6	7	8
	Italy			Spain				
	Colleague	Coauthor	Same subfield	Colleague	Coauthor	PhD Advisor	PhD committee	Research overlap
Female candidate	0.0026	0.0007	0.0209	-0.0012	-0.0003	-0.0001	-0.0010	0.0065
	(0.0004)	(0.0003)	(0.0060)	(0.0014)	(0.0002)	(0.0001)	(0.0003)	(0.0028)
Female evaluator	0.0017	-0.0015	-0.0067	0.0006	-0.0015	-0.0013	-0.0047	-0.0110
	(0.0009)	(0.0004)	(0.0075)	(0.0014)	(0.0002)	(0.0002)	(0.0006)	(0.0017)
Female candidate*	0.0029	0.0022	0.0133	0.0043	0.0010	0.0005	0.0013	0.0042
Female evaluator	(0.0007)	(0.0005)	(0.0045)	(0.0016)	(0.0002)	(0.0002)	(0.0005)	(0.0022)
Constant	0.0262	0.0140	0.5897	0.0453	0.0045	0.0025	0.0142	0.1959
	(0.0002)	(0.0001)	(0.0029)	(0.0007)	(0.0001)	(0.0000)	(0.0002)	(0.0010)
Observations	2,555,839	2,555,839	1,373,825	5,445,067	5,445,067	5,445,067	5,445,067	4,711,621

Notes: OLS estimates. The number of observations corresponds to the number of possible pairs between candidates and eligible evaluators with non-missing information in a given exam. In Italy, only evaluators who are based in an Italian university are considered. All regressions include exam fixed effects. Standard errors are clustered by candidate.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

- Networks are gendered:
  - Same affiliation: same-sex links are 13% more likely than mixed-gender links in Spain and 9% more likely in Italy
  - Co-authorship: same-sex links 22% more likely than mixed-gender links in Spain and 19% more likely in Italy
  - PhD supervisions: female candidates are 20% more likely to have a female advisor
- Networks matter for promotion (Zinovyeva and Bagues 2015, Bagues, Sylos-Labini and Zinovyeva 2015)
  - Colleague premium is 10% in Italy and 41% in Spain.
  - Co-author premium is 14% in Italy and 113% in Spain.
  - Advisor premium is 82% in Spain
- But remember that it is very rare to have a connection in the committee!
  - Colleagues (2.8% in Italy and 4.6% in Spain)
  - Coauthors (1.4% in Italy and 0.4% in Spain)
  - Student-advisor relationship (0.2% in Spain)

Table: **Connections and research similarity**

	1	2	3	4	5	6	7	8
	Italy				Spain			
Female candidate	0.008	0.006	-0.008	-0.010	-0.011	-0.010	-0.011	-0.011
	(0.008)	(0.007)	(0.009)	(0.009)	(0.007)	(0.007)	(0.008)	(0.008)
Female candidate * Share of female evaluators	-0.128	-0.124	-0.061	-0.060	-0.016	-0.020	-0.017	-0.021
	(0.035)	(0.035)	(0.046)	(0.046)	(0.028)	(0.028)	(0.035)	(0.035)
Connections in committee:								
Colleagues		0.181		0.180		0.319		0.319
		(0.036)		(0.044)		(0.031)		(0.031)
Coauthors		0.237		0.201		0.869		0.840
		(0.048)		(0.053)		(0.140)		(0.142)
PhD advisors						0.633		0.575
						(0.107)		(0.115)
PhD thesis committee						0.174		0.166
						(0.037)		(0.038)
Research similarity:								
Same subfield				0.046				
				(0.032)				
Overlap in research interests								0.124
								(0.037)
Controls:								
Expected connections		Yes		Yes		Yes		Yes
Expected same subfield				Yes				
Expected overlap in research interests								Yes
Number of observations	69020	69020	35832	35832	31243	31243	27998	27998

Gender segregation across subfields

- Gender segregation across subfields $\rightarrow$ same-sex preference
  - Same-subfield preference ✓
  - Gender segregation ×

Table: **Gender pattern of links between candidates and eligible evaluators**

	1	2	3	4	5	6	7	8
	Italy			Spain				
	Colleague	Coauthor	Same subfield	Colleague	Coauthor	PhD Advisor	PhD committee	Research overlap
Female candidate	0.0026	0.0007	0.0209	-0.0012	-0.0003	-0.0001	-0.0010	0.0065
	(0.0004)	(0.0003)	(0.0060)	(0.0014)	(0.0002)	(0.0001)	(0.0003)	(0.0028)
Female evaluator	0.0017	-0.0015	-0.0067	0.0006	-0.0015	-0.0013	-0.0047	-0.0110
	(0.0009)	(0.0004)	(0.0075)	(0.0014)	(0.0002)	(0.0002)	(0.0006)	(0.0017)
Female candidate*	0.0029	0.0022	0.0133	0.0043	0.0010	0.0005	0.0013	0.0042
Female evaluator	(0.0007)	(0.0005)	(0.0045)	(0.0016)	(0.0002)	(0.0002)	(0.0005)	(0.0022)
Constant	0.0262	0.0140	0.5897	0.0453	0.0045	0.0025	0.0142	0.1959
	(0.0002)	(0.0001)	(0.0029)	(0.0007)	(0.0001)	(0.0000)	(0.0002)	(0.0010)
Observations	2,555,839	2,555,839	1,373,825	5,445,067	5,445,067	5,445,067	5,445,067	4,711,621

Notes: OLS estimates. The number of observations corresponds to the number of possible pairs between candidates and eligible evaluators with non-missing information in a given exam. In Italy, only evaluators who are based in an Italian university are considered. All regressions include exam fixed effects. Standard errors are clustered by candidate.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table: **Connections and research similarity**

	1	2	3	4	5	6	7	8
	Italy				Spain			
Female candidate	0.008	0.006	-0.008	-0.010	-0.011	-0.010	-0.011	-0.011
	(0.008)	(0.007)	(0.009)	(0.009)	(0.007)	(0.007)	(0.008)	(0.008)
Female candidate * Share of female evaluators	-0.128	-0.124	-0.061	-0.060	-0.016	-0.020	-0.017	-0.021
	(0.035)	(0.035)	(0.046)	(0.046)	(0.028)	(0.028)	(0.035)	(0.035)
Connections in committee:								
Colleagues		0.181		0.180		0.319		0.319
		(0.036)		(0.044)		(0.031)		(0.031)
Coauthors		0.237		0.201		0.869		0.840
		(0.048)		(0.053)		(0.140)		(0.142)
PhD advisors						0.633		0.575
						(0.107)		(0.115)
PhD thesis committee						0.174		0.166
						(0.037)		(0.038)
Research similarity:								
Same subfield				0.046				
				(0.032)				
Overlap in research interests								0.124
								(0.037)
Controls:								
Expected connections		Yes		Yes		Yes		Yes
Expected same subfield				Yes				
Expected overlap in research interests								Yes
Number of observations	69020	69020	35832	35832	31243	31243	27998	27998

Stereotypes

- Do male evaluators hold stereotypes against women (unlike female evaluators)?
  1. Situations with evaluators are not familiar with candidates' research
  2. Sciences vs. Humanities and Social Sciences (Reuben, Sapienza and Zingales 2014)
  3. Less feminized fields
  4. Top positions

Table: **Heterogeneity analysis**

	1	2	3	4
	Italy		Spain	
Research overlap	$\geq median$	$< median$	$\geq median$	$< median$
	0.011	-0.179	0.081	-0.125
	(0.046)	(0.066)	(0.047)	(0.044)
Discipline	SSH	STEMM	SSH	STEMM
	-0.119	-0.133	-0.027	0.003
	(0.058)	(0.037)	(0.039)	(0.041)
Feminization of field	$\geq median$	$< median$	$\geq median$	$< median$
	-0.149	-0.072	-0.018	-0.016
	(0.042)	(0.057)	(0.040)	(0.037)
Level of promotion	FP	AP	FP	AP
	-0.111	-0.138	0.120	-0.072
	(0.059)	(0.038)	(0.054)	(0.032)

Notes: IV estimates. Standard errors are clustered by exam.

Outline

Summary of results

- Gender composition of committees does not affect significantly the quality of evaluations
- Female evaluators do not increase female success rates:
  - we can reject any positive impact in Italy
  - we can reject any sizable positive impact in Spain
- Interactions within committees are relevant
  - Women are slightly more favorable towards women...
  - ... but male evaluators are less favorable towards women in the presence of female evaluators.
- Gender does not play any role when evaluators belong to the same field of research as candidates $\Rightarrow$ focus on evaluators' knowledge

Policy implications

- No evidence in favor of gender quotas in the context of national evaluations
- Result might not necessarily hold in other contexts:
  - where fields are defined more broadly (and gender segregation is stronger)
  - where networks are more prominent (such as evaluations at the university level)
- Future work
  - Gender quotas: Spain 2007

Thank you for your attention!