Does random selection of commissioners improve the quality of selected candidates?

An investigation in the Italian Academia

DOCUMENTO DI VALUTAZIONE N. 3

DOCUMENTO
DI VALUTAZIONE

Ufficio Valutazione Impatto
Impact Assessment Office

Senato della Repubblica

Authors:

DANIELE CHECCHI, Fondazione Bruno Kessler - Istituto per la Ricerca Valutativa sulle Politiche Pubbliche (FBK-Irvapp)

SILVIA DE POLI, FBK-Irvapp

ENRICO RETTORE, FBK-Irvapp

We thank Nicolò Fraccaroli and Nicola Bazzoli for research assistance in data collection. We thank Giovanna Labartino for contributing to the first part of the project.

JEL CODES: D82, I23, J45, M51

KEYWORDS: UNIVERSITY RECRUITMENT, INCENTIVES, NEGOTIATION, FORMAL PROCEDURES

Does random selection of commissioners improve the quality of selected candidates?

An investigation in the Italian Academia

July, 2017

Abstract

We study a reform occurred in Italy in 2008 in the formation of selection committees for qualifying to university professorship. Prior to the reform members of the committees were elected by their peers, then they have been randomly drawn. This policy was intended to increase the equality of opportunities of candidates via a reduction of the role played by connection to commissioners. Results show that candidates internalised the changed environment and adapted their strategy of application. However the reform did not necessarily raise the impact of scientific quality of candidates on the outcome of competitions.

In questo lavoro studiamo gli effetti di una riforma delle regole di formazione delle commissioni di concorso a posizioni di professore universitario occorsa nel 2008. Prima della riforma le commissioni venivano elette dai docenti appartenenti all'ambito disciplinare del concorso; la riforma ha introdotto una selezione casuale dei commissari. La riforma si proponeva di dare pari opportunità a tutti i candidati al concorso riducendo l'importanza delle connessioni dirette con i commissari nella determinazione dell'esito. I risultati mostrano che i candidati hanno reagito alla riforma modificando le loro strategia di partecipazione ai concorsi. Ma la riforma non ha fatto aumentare l'importanza della qualità scientifica dei candidati nella determinazione dell'esito.

Contents

Introduction

The Italian academia is often described as plagued by nepotism and localism, which would be responsible for average poor performance as well as for the large degree of heterogeneity in research performance across university departments (Perotti, 2008, Durante et al. 2011, Moss 2012). This may explain why the formal rules of the hiring procedures have been repeatedly changed over the past decades, moving from national to local competitions in 1998 and back to a mixed system of national qualifications followed by local competitions in 2010. In the intermediate period, less radical changes were also introduced (reduction in the number of qualifications available in local competitions in 2003 and again in 2009). This paper focuses on one of these reforms, introduced in 2008, in a period in which competitions took place locally at each department: before the reform selection committees were elected by all academics in the field, while after the reform to a large degree the composition of each committee was randomly determined.

This reform was intended to reducing the advantage of local (often inbreeded) candidates, though the availability of a maximum of two qualifications per competition led very frequently to the assignment of one of them to a local applicant. By making the preordaining of competition outcomes less predictable, the reform also changed the behaviour of candidates, who were induced to expand the number of applications, though they had to withdraw a large fraction of them later on, given the existence of a limit of maximum five applications per year. Randomly selected committees may have dissimilar preferences about ranking candidates, and may be forced to rely more on objective measures of candidates' quality. Thus a second desirable outcome of the reform could have been the increased salience of scientific visibility (that we are going to proxy with the H index of the candidate) in the process of selection.

We take local competitions as our units of analysis, which we observe both before and after the reform. We obtain information on candidates and commissioners from the final report compiled at the conclusion of each competition. Given the burden of work required for data collection, we limit ourselves to four large Italian universities observed over the 2003-2008 time window, covering 664 competitions in all disciplines. Despite the final report being typically silent about the internal discussion among commissioners, we expect disagreement among randomly selected members being more frequent, and their influence in the process being possibly based on their own scientific visibility (also measured by their own H index).

Descriptive statistics suggest that candidates qualified after the reform have a higher scientific visibility than those obtaining the qualification before. One would be tempted to claim that randomized committees select better candidates, but unfortunately this exceeds what we are able to prove in our dataset. In facts we are unable to identify which channels made this possible (attraction of better candidates, more intense self-selection through withdrawal of application, better selection of commissioners, or a combination of them), due to lack of detailed information (which commissioner supported which candidate) and incomplete coverage of the process (we cover 11% of all competitions held in the country in the sample

period). Thus we limit ourselves to pointing out that randomized committees are effective in diluting the effect of pre-existing connections.

The paper is organized as follows. The next section describes the institutional change introduced by the 2008 reform, and contrasts its peculiarity against an emerging literature on randomized selection committees. Section 2 illustrates our dataset and section 3 presents our results, including some robustness checks. Section 4 summarises the findings and section 5 concludes.

1. The Italian recruitment of academics and the 2008 reform

The first decade of the 2000's has been characterised by a radical change in the selection of university professors in Italy. While in the 90's professors were selected in a national competition held every 4-5 years, in year 2000 competitions became local.[1] Conditional on availability of funding allocated by the Senate of each university, each department in need of filling in a position open a vacancy and form a selection committee. The committee could assign up to three qualifications ("idoneità", then reduced to two in 2003, due to an unexpected "inflation" of qualifications), among whom the department chose the preferred candidate. The other qualified candidates could then be hired by other universities (for a more detailed description see also Moss 2012)

The transition from a nationwide competition (where the number of selected candidates coincided with the number of open positions) to local ones have been criticized as promoting nepotism in the academia as well as for being detrimental for the quality of the research. Durante et al. (2011) show that the transition to the local competitions has been accompanied by an increase of homonymy in some university departments, which they interpret as evidence of familyism affecting the low social capital areas in the country. Alesina (2011) somehow complements their analysis, confirming the existence of a territorial divide between the north and the south of Italy in the extent of nepotism.

However, contrary to the expectations of previous authors, Checchi et al. (2014) do not find any significant impact of the reform on scientific productivity of Italian academics, irrespective of the outcome measure (number of papers, citations, cumulated impact factor). They suggest that the timing of competitions and the consolidated tradition of using objective measures of productivity in some research fields provides the relevant incentives, irrespective of how selections are undertaken.

These potentially contradictory results highlight the importance of focusing on committee formation, as one of the crucial dimensions of the hiring process. Under the national competition regime, selection committees resulted from a combination of elections (within each research field) and random draw.[2] That system was considered excessively cumbersome and time consuming. As a matter of fact, it resulted in a relatively large amount of time devoted

[1] The law (DPR n.390), approved in October 1998, began to take effect after the summer of 1999, so that the first promotions under the new rules took place towards the end of 1999.

[2] Committees were composed by an odd number of members (5, 7 or 9, depending on the expected number of applications, which are clearly correlated with the size of the relevant research field). In the case of (national) competition for associate professorships, the committee was made of 2, 3 or 4 associate professors and 3, 4 or 5 full professors, all belonging to the same research field where the vacancy is called for. In the competition for full professorships, all members have to be full professors affiliated to an Italian university. In order to keep control of the discipline, the formation of selecting committees for full professorships were exposed to less uncertainty: the constituency of all full professors in a specific research field (among the existing 371) elected a triple number of potential members, and the Ministry randomly drew the actual committee. Vice versa, while forming selecting committees for associate professorships, the Ministry drew a threefold number of potential members from each relevant constituency, and each component (full and associated professors) elected the committee.

to fine tune strategic alliances among different wings in the same discipline, in order to minimize the risk of having a "winner takes all" outcome in the formation of the selection committee. For this reason, when moving to the local competition regime it was substantially liberalized: each department opening a vacancy had the right to appoint one member (typically from the faculty), while the other four members were elected by the constituency represented by all professors in the field.[3] However this set-up often gave rise to draw-backs. Since being elected as committee member was not directly rewarded, university professors had a sort of "gift exchange" as the unique incentive to become members of the selection committee. Academics accepted to become candidates in the electoral process leading to the committee formation in exchange of a promise of future reciprocation from colleagues in departments requesting their participation. While these incentives were appealing to pro-social attitudes, other "illicit" exchanges were also available: a professor appointed as internal member could invite other colleagues from other departments with the promise of assigning one qualification to their preferred candidates.

This mal-practice of preordaining the outcome of a competition before its start was clearly unfair to almost all participants in the competition, and was early lamented in press and in books (Perotti 2008). Different Ministries of Education tried to limit the incentives to malpractices, until the year 2010 when the system was again radically reformed, moving back to a national examination to obtain a qualification then followed by local competitions restricted to qualified candidates.[4] One of these attempts prior to the 2010 reform – currently in operation - is the reform we examine in this paper: in 2008 the Ministry of Education issued an apparently marginal reform in the procedures of forming the selection committees.[5]

Replicating the selection principles once prevailing in national competitions, the constituency of academics in a research field was required to elect twelve potential members, among which the Ministry would have drawn four actual members. If the number of full professors in a given research field was lower than 12, all full professors were included in the pool, supplemented in case of need by professors of related disciplines, and random selection was then applied.[6] The introduction of randomness in the formation of the committee was intended to preclude the possibility of pre-agreement, because the incentive of pre-assigning

[3] While the size of the committee was fixed to five for all research fields, limitation in the number of candidates per competition was introduced by imposing a limit of a maximum number of applications to five per candidate each year.

[4] See the law n.240/2010. For a description of the new system see Bagues et al. 2014.

[5] See the article n.1 of DL n.180 issued on 10/11/2008, n. 180 ("Disposizioni urgenti per il diritto allo studio, la valorizzazione del merito e la qualità del sistema universitario e della ricerca"), then converted into the law n.1/2009 dated 9/1/2009. The procedures were then specified by the Decreto Ministeriale n.139 dated 27/3/2009 ("Modalità di svolgimento delle elezioni per la costituzione delle commissioni giudicatrici di valutazione comparativa per il reclutamento dei professori e dei ricercatori universitari").

[6] In 2008 there were 74 research fields (settori scientifico-disciplinari) out of 371 with a number of full professors less or equal to 13, which therefore were exposed to pure randomization (since one professor has to serve as internal member and twelve consisted as the pool for selection). This however represents a lower bound, because each professor could serve as commissioner only once a year: thus two competitions in the same year and in the same research field would have required more than 26 professors in order to avoid pure randomization (and so on).

the qualifications was diluted among a large number of participants. In the sequel, we find clear evidence that this outcome has been attained, because the existence of connections between candidates and commissioners loses impact on the probability of being selected.

The decline in the probability of facing pre-existing connections between commissioners and candidates attracted additional applications, from candidates who would have been otherwise discouraged. However, as the composition of the commission became revealed, some of these applications were withdrawn, due to the limit imposed on the maximum number (five) of applications per year.

In principle, pre-existing connections are not always detrimental, since they provide information about otherwise unknown candidates. With the decline in the probability of candidates-commissioners connections, one would have expected other source of information becoming more relevant. Thus we would have expected scientific productivity and visibility, objectively inferred from the applicants' CV, becoming more relevant among the selection criteria. However, different commissioners may have different opinions about the importance of publishing papers in journals, often in relationship with the research field tradition and their own experience within it. Thus we are not surprised by the empirical analysis where the measure of scientific visibility (proxied by the H index of the candidates) increases its impact only in a subset of research fields (the so-called bibliometric sectors).

These expectations are in line with a related literature on academic competitions. In the literature random selection of examiners is typically exploited to support the causal interpretation of the correlation between observable characteristics of commissioners and the outcome of selections. In a paper exploiting the same reform we focus on De Paola and Scoppa (2015) study the impact of evaluators' gender on the probability of appointment to professorship of female candidates. Their sample consists of 130 competitions held in two research fields (economics and chemistry) in the aftermath of the reform (vacancies opened in 2008). They provide robust evidence that gender composition of selection committees matters, and that female commissioners reduce gender discrimination experienced by female candidates.

Zinovyeva and Bagues (2015) exploit random matching of candidates to evaluators to measure the (causal) effect of pre-existing connections on the probability of being qualified as professors in Spain during the period 2002-2006. Due to the lack of rules on conflict of interests it happens that candidates are evaluated by professors who have been their PhD supervisors, or co-authors, or even colleagues; they find that candidates are 50 percent more likely to be promoted when the committee includes commissioners who are strongly connected to them. They consider the positive contribution of connections as source of information, distinguishing between good and bad view of networks; however their conclusion is that biased assessments prevail, identified by looking at scientific productivity in the following five years after selection.[7]

[7] Similar approach is contained in Godechot (2015): shortlists for faculty recruitment at one leading French university (the École des Hautes Études en Sciences Sociales – EHESS) between 1960 and 2005 were compiled by a randomly selected committee made of department members. Thanks to information on PhD supervisorships, the author is able to measure the gain in probability experienced by a candidate

Random matching of selectors and candidates is also exploited in the analysis of the recently introduced qualification round for Italy (Bagues et al. 2014 and 2015). In the first paper the authors find that having women in the examining committees is detrimental to the chance of female candidates to be promoted,[8] concluding that the introduction of gender quotas in evaluation committees might have unintended consequences. In the second paper they examine the role of connections (colleagues and/or co-authors) in the probability of withdrawing from competitions, showing that weak candidates are more likely to withdraw their applications once a connected examiner is randomly drawn. Differently from Zinovyeva and Bagues (2015), they argue that self-selection of applicants due to application withdrawals is consistent both with a pro-bias and with a pro-information interpretation of the role of academic connections.[9] This self-selection is also gendered, as shown by DePaola et al. (2014), where they consider as potential applicants all assistant or associate professors in the Italian academia, and they show that the gender gap in the probability of applying for a competition holds true only for individuals in the lower tail of the distribution of scientific productivity.

With respect to previous literature, we take a different perspective and we analyse the impact of introducing randomisation in the commission formation. Randomisation was intended to increase equality of opportunities among candidates, against the bad habit of nepotism and/or simply inbreeding. We show that this outcome was achieved via reduction of the impact of connections between candidates and commissioners (here measured by common affiliation between commissioners and candidates). However, previous literature stresses the double-face nature of connections, because they may bias evaluators' judgments, but they may also reduce informational asymmetries about otherwise unknown candidates. As a consequence, we would have expected an increase in the explanatory power of observables (like scientific productivity), which is not always observed in the data. The competition outcome is analysed as a two-stage procedure, for candidates had the possibility of withdraw their applications before the selection process starts . Since we do not cover all competitions taking place during our sample period, we cannot investigate the drivers of candidate applications, and therefore we cannot assess the actual degree of self-selection of candidates. Still, we do observe how they react as the information on the composition of the recruiting committee become publicly available.

---

whose supervisor was randomly selected against an otherwise identical candidate whose supervisor was not selected.

[8] "Each additional female evaluator decreases by two percentage points the success rate of female applicants, relative to male applicants" (Bagues et al. 2014, p.5)

[9] "Researchers in the top tercile in terms of their research output are 5 p.p. more likely to succeed when the committee includes a co-author or a colleague. Weaker researchers also benefit from connections by not making costly errors in application decisions. Researchers in the bottom tercile are 6 p.p. less likely to apply when the evaluation committee includes a co-author or a colleague and their chances of success are 3 p.p. higher. As a result, the probability that they fail the evaluation is 9 p.p. lower. Evidence from a subsequent round of evaluations suggests that, by postponing their application, weak researchers with a connection in the committee benefit also from higher success rates in the future. Overall, the evidence is consistent with the existence of a bias in favour of connected candidates and also with the notion that connections reduce information asymmetries." (Bagues et al. 2015).

È possibile utilizzare queste raccolte per inserire tabelle, intestazioni, piè di pagina, elenchi, frontespizi e altri blocchi predefiniti per i documenti. Anche le immagini, i grafici o i diagrammi che vengono creati sono coordinati con l'aspetto del documento. È possibile modificare rapidamente la formattazione del testo selezionato nel documento scegliendo uno stile veloce dalla raccolta stili veloci disponibile nella scheda Home.

È inoltre possibile formattare il testo direttamente utilizzando gli altri controlli della scheda Home. Con la maggior parte dei controlli è possibile scegliere di utilizzare l'aspetto del tema corrente oppure un formato specificato direttamente dall'utente. Per cambiare l'aspetto generale del documento, scegliere un nuovo tema nella scheda Layout di pagina. Per modificare gli stili disponibili nella raccolta stili veloci, utilizzare il comando Cambia stili.

2. Our dataset

We use data from different sources. We collected data from the final report (verbale riassuntivo) produced by each selecting committee in competitions to associate and full professor positions opened in four large Italian universities - Padua, Milan 'Statale', Naples 'Federico II' and Rome 'La Sapienza' - from 2003 to 2008.[10] For each competition, we have information about the recruiting committee, the pool of applicants, the candidates that withdrew from the competition before its conclusion (including those who did not show up for the interview)and the winners.[11]

The competitions included in our sample qualifies (up to) two candidates, among which the local department chooses the one to be appointed. The department can also refuse to hire any one in case they do not like the qualified candidates. In this case the department cannot open a vacancy for thei same position over the subsequent 12 months. The qualified candidate(s) not recruited by the department opening the position can be hired by any other Italian department.

Candidates were allowed to apply to a maximum of five competitions per year both before and after the reform. Anecdotal evidence suggests that some candidates strategically applied to more than five competitions per year, withdrawing excess applications once the composition of the selecting committees was revealed. However this is a risky strategy because if the

[10] By the end of the sample period (31/12/2008), these universities accounted for almost 20% of the wholeItalian academia: their faculties (including tenured full, associate and assistant professors) were 2427 (Milan 1st university, known as Statale), 2382 (Padua), 3017 (Naples 1st university, known as Federico II) and 4704 (Rome 1st university, known as La Sapienza), against an overall population of 63290 academics. In terms of quality, the most recent research assessment exercise (VQR 2004-2010 – table 6.10a of the final report,) ranked the Italian top five universities in the following order: Bologna, Rome, Padua, Milan and Naples.

[11] In some research field it is a common (still undocumented) practice to let the candidates know they have limited chance to be selected in a specific competition. Since each candidate could not apply to more than 5 positions per year one possible outcome of this communication is the withdrawal of the application.

withdrawal of the application was received after the first meeting of the committee, the candidate was excluded from all competitions s/he has applied for. An alternative interpretation applying to the "before reform" competitions is that once the committee is formed and shortlist of candidates were informally agreed upon (even before the first meeting), some commissioner may have actively pressed out-of-list candidates to withdraw from of the competition, under the threat of a negative judgment in the official report. Bagues et al. 2015 obtain similar results for national qualification exams, and they interpret it as evidence of information provision associated to academic ties.

Candidates and members of selecting committees

For each candidate and member of committees we have information about age, gender, the university of affiliation at the time of competition and their research output. Age is obtained from administrative archives for academics already tenured, or searching CVs on-line otherwise.[12] The affiliation of candidates and evaluators are obtained by open data sources of the Ministry of Education and University-Miur (see http://cercauniversita.cineca.it/php5/docenti/cerca.php).

The research output is obtained from two alternative data sources, Publish or Perish (PoP), which is based on Google Scholar, and Scopus (www.scopus.com), promoted by Elsevier. The advantage in using Scopus is the reduction in the risk of homonymous confounders, since each author can be identified with affiliation and research field. On the other hand, Scopus does not properly cover humanities, where scholars often write in Italian; in this case, we use PoP to collect data on these fields of study.

For each applicant and commissioner, we have collected the number of papers, the number of citations and the H index (namely the number of papers that receive at least an equivalent number of citations). Note that this information on the research standing of candidates and commissioners are referenced to the date we extracted them from the web (from March to December 2014) because it would have been an excessive workload to obtain this information referred to the time at which competitions took place, especially for the PoP database. However, to understand whether this approximation may bias our analysis, we collected the research standing for a small subsample of candidates over the period 2003-2014. The comparison between the H index as of the date of extraction and the corresponding measure at the time of the competition suggest that by this approximation we do not introduce systematic measurement errors in the analysis.[13]

Since the features of the distribution of the number of publications and of the H index vary significantly across fields of study, we standardise them using the field specific average and standard deviation as computed on candidates included in our sample who applied

[12] Whenever we do not find the date of birth we use the year of graduation to estimate the age; when this is absent, missing values are replaced them with the average age of candidates in the field of study.

[13] This analysis is discussed in section A of the Appendix 1.

before the reform.[14] Candidates who applied after the reform are not included in this calculations to avoid a potential bias that could have been created by the reform itself.

Competitions

There are 664 competitions included in the sample, of which 230 are for full professorships and 434 for associate professor positions[15]. Table 1 shows the average number of candidates and the fraction of withdrawals in competitions before and after the reform. Descriptive statistics are presented separately for two groups of research fields: the bibliometric disciplines (named this way because of the prevailing pattern of research assessment based on bibliometric indices), include mathematics and informatics, physics, chemistry, earth sciences, biology, medicine, agricultural and veterinary sciences, engineering, psychology; all the remaining, non-bibliometric disciplines, are humanities, literature, history, philosophy, art history, pedagogy, law studies, economics and statistics, political and social sciences, architecture. Table 1 shows that prima facie the reform resulted in a large increase of the number of candidates. This is likely due to the increased uncertainty about the committee composition to which potential candidates reacted by submitting more applications[16].

The percentage of withdrawals is higher in competitions for associate professors positions (in the order of half of the applicants) since participation costs are higher: the candidate has to be personally interviewed at the university opening the position[17].

Selecting committees and candidates

Table 1: Descriptive statistics of competitions for professorship, years 2003 to 2008, University of Padua, Milan – *Statale*, Rome – *La Sapienza*, Naples – *Federico II*.

	Before the reform			After the reform		
	All	Non bibliometric	Bibliometric	All	Non bibliometric	Bibliometric
			Full professors			
Average number of candidates	8.8	5.8	10.0	16.1	13.6	16.9
% of withdrawals	20.1	17.2	21.2	11.3	12.2	11.0
Number of competitions	138	38	100	92	22	70
			Associate professors			
Average number of candidates	9.6	7.7	10.3	18.7	14.8	20.1
% of withdrawals	50.0	40.8	53.3	47.4	39.3	50.4
Number of competitions	273	72	201	161	43	118

[14] In section A4 of the Appendix 1 we report the number of observations, the mean and the standard deviation used for this standardisation (see tables A4 and A5).

[15] 9% of the competitions are excluded from the analysis because we could not recover their official final reports.

[16] An alternative reason for this increase is related to irregular time distribution of competitions over years, since only 6% of competitions took place in 2006, while no competition is observed in 2007 in our sample.

[17] The recruitment of full professors did not require any interview nor any teaching test if the applicant already held a position as associate professor.

Table 2 provides some descriptive statistics on selecting committees. After the reform in competitions for both associate and full positions there were more than half of committees with at least one woman, while before the reform this percentage was significantly lower (40.6%), at least for competitions for full professorship. The comparison between the H index of the commissioners before and after the reform is not straightforward to interpret, since the difference between the H index collected in 2014 and the H index collected at the time of the competition increases over time (see section A in the Appendix 1). At face value, statistics reported in table 2 suggest that the reform increased the average research quality of evaluators as well as their variability in competitions for associate professorships.[18] It can be noticed that also the H index of the internal commissioner appointed by the departments slightly increased. As a possible reaction to the reform, departments started to appoint internal examiners with a better research record, both in the case of competitions for associate and full professorship. The overall effect is an increase in the quality of the selecting committees.

The random selection of external members should prevent possible pre-agreements between the internal member and the supposed-to-be-elected commissioners, thus reducing the "similarity" among them. If we take the correlation index between their H indices as a proxy for such a similarity, we observe a decline in the case of competitions for associate professorships (which is even more surprising, while considering the substitution of associate professor members with full professor ones), but an opposite trend in the case of full professorships.

[18] The increase in research quality of evaluators in competitions for associate professorship is partially a direct outcome of the reform, since it abolished the presence of two associate professors in the committee, leading to committees composed by full professors only. Also the increased variance has to be considered in this perspective.

Table 2: Descriptive statistics of selecting committees, years 2003 to 2008, University of Padua, Milan – *Statale*, Rome – *La Sapienza*, Naples – *Federico II

	Before the reform			After the reform		
	All	Non bibliometric	Bibliometric	All	Non bibliometric	Bibliometric
	competitions for full professorship					
fraction of selecting committees with at least one woman	40.6	63.2	32	53.3	54.5	52.9
average H index of the internal commissioner	0.12	0.56	-0.04	0.17	-0.04	0.23
average H index of the external commissioners	0.05	0.49	-0.12	0.03	0.31	-0.06
Standard deviation of the H index of the external commissioners	0.98	1.37	0.83	0.90	0.92	0.89
Correlation between the H index of the internal and external commissioners	0.16	0.01	0.20	0.22	0.03	0.32
Average distance of external evaluators from the university posting the vacancy (Kms measured by Google maps)	454	415	469	519	509	522
	competitions for associate professorship					
fraction of selecting committees with at least one woman	59.3	72.2	54.7	54.7	81.4	44.9
average H index of the internal commissioner	0.56	0.72	0.50	0.71	0.60	0.75
average H index of the external commissioners	0.18	0.40	0.11	0.51	0.71	0.43
Standard deviation of the H index of the external commissioners	0.98	1.06	0.96	1.19	1.54	1.06
Correlation between the H index of the internal and external commissioners	0.26	0.16	0.30	0.14	0.13	0.22
Average distance of external evaluators from the university posting the vacancy (Km's measured on Google maps)	434	417	440	517	501	523

The H index is standardised within each scientific field

Another effect of the reform is the moderate increase of the average distance between the university of affiliation of external members and the location of the competition. After the reform, the distance is increased by approximately 80 km in associate professor competitions and by 65 km in full professor competitions. This confirms the intuition that when external commissioners were elected there was a self-selection of near-by evaluators, who were more likely to be connected to the departments posting the vacancies.

Table 3 reports descriptive statistics for all applicants (columns 1 and 4), excluding candidates who eventually withdrew from the competition (columns 2 and 5) and for candidates who were eventually selected as qualified (columns 3 and 6). Candidates applying in competitions for full (associate) professorships are on average 49 (43) year old and the share of female applicants is about 27% (37%). These figures do not change restricting our attention on candidates who decided not to withdraw and on winners.

As regards the affiliation, before the reform, 23.9% of applicants for full professorships (22.6% for associate positions) were insiders,namely they were applying for vacancies posted by their department, while after the reform it declined to 17.2% and 18.7%, respectively. The percentage of insiders is slightly higher among those not withdrawing their application, in particular in competitions to associate professorship. Finally, it is much higher as well as unaffected by the reform among winners of competition. Note that a fraction of insiders

close to 1/2 among winners suggests that one insider was selected among the two qualified nearly always (and she is also likely to having been hired by the department opening that vacancy). The average H index among all applicants before the reform is zero by construction (since standardisation is based on the pool of pre-reform applicants). It is slightly higher among qualified candidates and the applicants-winners differential is larger after the reform. While we cannot take this as direct outcome of the reform, it is suggestive of an improvement as a result of randomly selecting the external commissioners.

Tabella 3: Descriptive statistics of candidates, years 2003 to 2008, University of Padua, Milan – *Statale*, Rome – *La Sapienza*, Naples – *Federico II*

	Before the reform			After the reform		
	All	Net of withdrawals	Qualified	All	Net of withdrawals	Qualified
			Full professors			
Share of female	27.4	27.5	27.2	26.8	27.2	28.1
Age	48.3	48.7	48.8	50.0	50.0	49.5
Share of insiders	23.9	26.7	54.1	17.2	18.2	53.4
H index (standardised)	0.00	-0.05	0.12	0.05	0.04	0.30
Total	1217	975	246	1480	1325	178
			Associate professors			
Share of female	37.2	36.8	36.0	38.1	35.2	37.3
Age	42.5	43.1	43.6	43.0	43.1	43.8
Share of insiders	22.6	34.4	50.2	18.7	27.1	49.4
H index (standardised)	0.00	0.01	0.11	-0.01	0.03	0.23
Total	2619	1152	506	3010	1541	308

3. Analysis

To study the determinants of the probability of success among candidates - and the way it changed after the reform - we estimate a linear probability model with competitions fixed effect. This way we control for unobservable systematic differences across disciplinary fields, locations and timing of the competitions. Thus, coefficients of the explanatory variables are identified exploiting their within competition variability, i.e. the differences between candidates taking part in the same competition.

We measure the effect of the applicant productivity (proxied by the standardized H index), age, gender and insider status, i.e. whether the applicant holds a position at the department posting the vacancy[19]. To account for connections between candidates and commissioners, we also include a dummy variable for candidates ffiliated to the same university of an external commissioner.

Eventually, we add further interaction terms between the commissioners' and the candidates' characteristics. We introduce a dummy taking value one when a female applicant is matched to a selecting committee composed only by men (DePaola and Scoppa 2015). We also include an interaction between the scientific productivity of the external commissioners relative to the internal one and the scientific productivity of the candidate. It is a

[19] By holding a position we mean that the applicant is an academic in the payroll of the university.

dummy variable equal to one when the external commissioners are on average better than the internal one with respect to the H index multiplied by the H index of the candidate. This is to capture possible reversal in the bargaining power of external vs internal commissioners driven by their scientific reputation.

Table 4 shows the estimates of the model for competitions to associate professorship. Results are presented separately for the two periods before and after the reform and for bibliometric and non-bibliometric fields. Regressions are estimated on the whole pool of applicants including those withdrawing from the competition.

The most important individual characteristics for the probability of obtaining the qualification is by far a connection to the commissioners. Before the reform being an insider increased the chances of success by 0.35 in bibliometric disciplines and 0.40 in the non-bibliometric ones, while having a commissioner from the same university increased the probability of success by 0.19 and 0.29, respectively.

After the reform the effect of being an insider is slightly smaller both n bibliometric and in non bibliometric disciplines while the effect of having a commissioner from the same university drops quite a lot, from 0.19to 0.04 and from 0.29 to 0.19, respectively. This is likely due to the random selection of external commissioners which made less easy an ex ante agreement among commissioners.

As for the role of the candidate scientific productivity, in line with expectations its effect is positive and statistically significant, but this effect is quite small in magnitude, both before and after the reform. One standard deviation increase in the H index in bibliometric disciplines yields an effect as large as 0.05 and 0.04 before and after the reform, respectively. It is even smaller in non bibliometric disciplines.

Adding the interaction between candidates and committee (regressions in even columns), does not produce any remarkable difference. Gender has no significant impact on the probability of success, also after controlling for the committee gender composition.

Tabella 4: Fixed effect linear probability model for the probability of winning a competition to associate professorship, years 2003 to 2008, University of Padua, Milan – *Statale*, Rome – *La Sapienza*, Naples – *Federico II*

	Bibliometric				Non Bibliometric			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
VARIABLES	Before	Before	After	After	Before	Before	After	After
H index	0.047***	0.056***	0.042***	0.046***	0.040**	0.035	0.014	0.030
	(0.010)	(0.013)	(0.009)	(0.012)	(0.017)	(0.035)	(0.013)	(0.023)
Age	-0.021	-0.021	0.009	0.009	-0.012	-0.014	0.008	0.009
	(0.016)	(0.016)	(0.009)	(0.009)	(0.024)	(0.024)	(0.017)	(0.018)
Age²/1000	0.269	0.276	-0.095	-0.094	0.125	0.149	-0.097	-0.102
	(0.178)	(0.178)	(0.100)	(0.101)	(0.264)	(0.265)	(0.201)	(0.203)
Female	0.002	-0.000	-0.009	-0.016	-0.054	-0.073	-0.026	-0.020
	(0.020)	(0.023)	(0.012)	(0.015)	(0.036)	(0.047)	(0.030)	(0.034)
Female×No female evaluators		0.006		0.015		0.060		-0.025
		(0.042)		(0.026)		(0.074)		(0.069)
External above internal×H index candidate		-0.025		-0.008		0.007		-0.023
		(0.020)		(0.018)		(0.039)		(0.027)
Insider	0.346***	0.345***	0.220***	0.219***	0.395***	0.394***	0.330***	0.329***
	(0.026)	(0.026)	(0.023)	(0.023)	(0.059)	(0.059)	(0.057)	(0.057)
Commissioner of the same university	0.192***	0.193***	0.036*	0.036*	0.292***	0.294***	0.191***	0.192***
	(0.030)	(0.030)	(0.022)	(0.022)	(0.055)	(0.054)	(0.059)	(0.060)
Constant	0.461	0.472	-0.148	-0.147	0.414	0.460	-0.098	-0.104
	(0.337)	(0.337)	(0.201)	(0.202)	(0.515)	(0.516)	(0.376)	(0.379)
Observations	2,062	2,062	2,374	2,374	557	557	636	636
R^2	0.144	0.144	0.093	0.094	0.174	0.175	0.130	0.132
Number of competitions	201	201	118	118	72	72	43	43

*Eteroskedasticity robust standard errors in parentheses, clustered at competition level, competition fixed effect – statistical significance *** p<0.01, ** p<0.05, * p<0.1*

Table 5 presents the results for the decision to withdraw. First note that candidates seem to have rationally anticipated the outcome of the competition: the very same factors that affect the likelihood of winning also affect with a reversed sign the likelihood of withdrawing. In particular, candidates who were outsider and/or did not have connections to the committee were more likely to withdraw. The reform seems to alter only the effect of being an outsider with a link to a commissioner: this effect is negative and statistically significant both before and after the reform but much smaller after the reform at least for bibliometric disciplines.

Consistently with the results from the regression for the probability of winning (Table 4), the key result here is that the effect of being an insider is nearly unaffected by the reform. That is switching from elected to randomly chosen external commissioners does not alter the advantage of being an insider (Table 4) and this advantage is rationally anticipated in the withdrawal decision (Table 5).

Looking at the research productivity, we observe that better candidates have a lower probability to withdraw.

Table 5: Fixed effect linear probability model for the probability of withdrawing from a competition to associate professorship, years 2003 to 2008, University of Padua, Milan – *Statale*, Rome – *La Sapienza*, Naples – *Federico II*

	Bibliometric		Non Bibliometric	
	(1)	(2)	(3)	(4)
VARIABLES	Before	After	Before	After
H index	-0.048***	-0.018	-0.012	0.029
	(0.017)	(0.019)	(0.040)	(0.035)
Age	0.004	-0.064***	-0.013	-0.066***
	(0.018)	(0.017)	(0.026)	(0.023)
Age²/1000	-0.094	0.701***	0.130	0.744***
	(0.203)	(0.186)	(0.284)	(0.262)
Female	-0.003	0.075***	0.044	0.059
	(0.032)	(0.027)	(0.053)	(0.053)
Female×No female evaluators	-0.026	-0.000	-0.045	-0.064
	(0.054)	(0.043)	(0.102)	(0.105)
External above internal×H index candidate	0.012	-0.044*	0.002	-0.038
	(0.025)	(0.026)	(0.051)	(0.041)
Insider	-0.333***	-0.326***	-0.300***	-0.284***
	(0.026)	(0.025)	(0.055)	(0.046)
Commissioner from the same university	-0.189***	-0.066*	-0.177**	-0.166**
	(0.036)	(0.034)	(0.068)	(0.068)
Constant	0.704*	1.959***	0.862	1.910***
	(0.399)	(0.382)	(0.581)	(0.520)
Observations	2,026	2,374	557	636
R^2	0.092	0.080	0.069	0.063
Number of competitions	198	118	72	43

*Heteroskedasticity robust standard errors in parentheses, clustered at competition level, competition fixed effect – statistical significance *** p<0.01, ** p<0.05, * p<0.1 - Note: three competitions were excluded because the final report was incomplete and*

we could not observe the name of candidates who withdrew

After the reform age becomes relevant for the decision to withdraw both in bibliometric and in non bibliometric disciplines. The probability of withdrawal decreases with age up to 45.

Competitions to full professorship

In Table 6 we estimate the fixed effect linear probability model for the probability to qualify as a full professor. As in the case of competitions to associate professorship, connections to commissioners play the most relevant role. before the reform, local candidates have an advantage as large as 0.38 of being selected in bibliometric research areas while it is 0.52 in non-bibliometric ones. After the reform the advantage is nearly the same in bibliometric disciplines while it declines to 0.32, in non bibliometric ones.

The role of links to a commissioner from the same university reacts to the reform similarly to the case of competition to associate professorships: it becomes much less relevant after the reform.

It is interesting to note that before the reform the coefficients associated to scientific productivity had similar magnitudes in competitions for full and associate professorships, while after the reform only selection for full professors in bibliometric sector records an increase in size: after the reform a one standard deviation increase of the H index rises the likelihood of being selected by 0.07. However, as in the case of competitions to associate professorship this is definitely not a huge effect (a one standard deviation increase amounts to 12 points in medicine or 5 points in mathematics).

Table 7 presents the results for the probability of withdrawal in full professorship competitions. The most relevant role is again played by the insider status, even if this effect is lower than in the case of associate competitions. After the reform, this effect is approximately halved. All the other observable characteristics do not get a statistically significant estimate.

Table 6: Fixed effect linear probability model for the probability of winning a competition to full professorship, years 2003 to 2008, University of Padua, Milan – *Statale*, Rome – *La Sapienza*, Naples – *Federico II*

	Bibliometric				Non Bibliometric			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
VARIABLES	Before	Before	After	After	Before	Before	After	After
H index	0.033**	0.027	0.073***	0.082***	0.024	-0.056*	-0.006	0.019
	(0.016)	(0.023)	(0.012)	(0.016)	(0.031)	(0.031)	(0.012)	(0.025)
Age	0.008	0.008	0.012	0.012	-0.070	-0.069	0.062*	0.063*
	(0.016)	(0.016)	(0.013)	(0.013)	(0.055)	(0.054)	(0.032)	(0.032)
Age²/1000	-0.089	-0.094	-0.132	-0.132	0.641	0.626	-0.651**	-0.664**
	(0.167)	(0.169)	(0.124)	(0.123)	(0.542)	(0.529)	(0.292)	(0.287)
Female	-0.019	-0.064	0.008	0.010	-0.014	0.006	-0.027	0.000
	(0.026)	(0.044)	(0.024)	(0.032)	(0.072)	(0.071)	(0.045)	(0.053)
Female×No female evaluators		0.066		-0.002		0.008		-0.085
		(0.055)		(0.047)		(0.157)		(0.083)
External above internal×H index candidate		0.011		-0.018		0.156***		-0.036
		(0.031)		(0.023)		(0.049)		(0.031)
Insider	0.383***	0.383***	0.339***	0.339***	0.520***	0.522***	0.324***	0.331***
	(0.037)	(0.037)	(0.036)	(0.036)	(0.076)	(0.078)	(0.048)	(0.049)
Commissioner from the same university	0.147***	0.148***	0.062**	0.063**	0.189*	0.197**	0.085	0.094
	(0.031)	(0.031)	(0.031)	(0.031)	(0.103)	(0.096)	(0.065)	(0.065)
Constant	-0.095	-0.110	-0.228	-0.228	2.006	1.990	-1.330	-1.347
	(0.391)	(0.395)	(0.337)	(0.333)	(1.381)	(1.346)	(0.877)	(0.868)
Observations	995	995	1,180	1,180	222	222	300	300
R^2	0.164	0.165	0.182	0.183	0.236	0.265	0.133	0.139
Number of competitions	100	100	70	70	38	38	22	22

*Heteroskedasticity robust standard errors in parentheses, clustered at competition level, competition fixed effect) – statistical significance *** p<0.01, ** p<0.05, * p<0.1*

Table 7: Fixed effect linear probability model for the probability of withdrawing from a competition to full professorship, years 2003 to 2008, University of Padua, Milan – *Statale*, Rome – *La Sapienza*, Naples – *Federico II*

	Bibliometric		Non Bibliometric	
VARIABLES	Before	After	Before	After
H index	0.015	0.021	0.095*	0.002
	(0.020)	(0.015)	(0.047)	(0.019)
Age	-0.046*	0.004	0.034	-0.068*
	(0.025)	(0.015)	(0.042)	(0.036)
Age²/1000	0.366	-0.043	-0.358	0.629*
	(0.261)	(0.142)	(0.405)	(0.336)
Female	0.015	-0.020	-0.008	-0.015
	(0.054)	(0.031)	(0.048)	(0.040)
Female×No female evaluators	-0.020	0.015	0.079	0.001
	(0.065)	(0.045)	(0.116)	(0.077)
External above internal×H index candidate	-0.012	0.025	-0.089	-0.016
	(0.027)	(0.021)	(0.061)	(0.022)
Insider	-0.103***	-0.047**	-0.178***	-0.072*
	(0.031)	(0.022)	(0.045)	(0.040)
Commissioner from the same university	-0.067*	0.016	-0.052	0.039
	(0.036)	(0.037)	(0.060)	(0.053)
Constant	1.603***	-0.001	-0.556	1.947**
	(0.590)	(0.392)	(1.058)	(0.936)
Observations	995	1,180	222	300
R^2	0.059	0.013	0.097	0.030
Number of competitions	100	70	38	22

*Heteroskedasticity robust standard errors in parentheses, clustered at competition level, competition fixed effect – statistical significance *** p<0.01, ** p<0.05, * p<0.1*

Robustness checks

Here we discuss some possible sources of bias in our results. First, there might be measurement errors attributable to the imperfect identification of candidates due to the presence of homonymous. However when excluding homonymous names we do not find significant differences in our results (see Table A1 of Appendix 1 and compare it to columns 6 and 8 of table 4:the statistical significance of the variables of interest do not change, although the magnitude of some coefficients do vary).

A second source of concern is the imperfect randomisation among the potential members of the committee, which were randomly drawn from an elected pool as large as three times the number of commissioners to be appointed. Some research fields (settore scientifico-disciplinare) did not have at least 13 full professors (one to be appointed as internal member and 12 in the pool from which the Ministry drew the final four), in which case no election too place, all professors were included in the pool and they were supplemented by professors belonging to related research fields in order to reach the required number of 12. We distinguish between the two cases - election+randomisation and pure randomisation - in tables A6 and A7 of Appendix 1. Estimates change a bit in the two cases with stronger effects of having connections in the subset of pure randomisation. This can be taken as an evidence of attenu-

ation of these effects due to to imperfect randomisation.

A final point regards the external validity of our results. In tables A8 and A9 we provide the distribution of the competitions by research area in our sample and in the entire Italian academia, finding that our sample over-represents the area of medical schools and underrepresents the schools of Economics and the schools of Law. Tables A10 and A11 display the time profile of the competitions and the (average) number of applicants: if we exclude some evidence of anticipation of competitions in 2004, we may conclude that our research strategy of selecting a subset of universities does not reduce the representativeness of our sample, reinforcing the external validity of our results.

4. Summing up

We have shown that changing the way recruiting committees are selected does affect the dynamics of selecting candidates qualified for professorship in four largest Italian universities. In the pre-reform regime the recruiting committees were elected, the four external members belonged to geographically closer universities, had a lower scientific visibility (as measured by the standardised H index, also as a result of the presence of two associate professors among the four). The scientific standing of the internal member increased a bit after the reform and the random selection has led to better qualified but more heterogeneous committees in competition for associate professorship. In addition, the fraction of competitions for full professorship with at least one woman in the committee rose to one half.

If the reform makes it harder to preordain the outcome by affecting the composition of the selecting committee, the reaction of candidates internalises the increased uncertainty of the reformed competitions: the number of applicants doubles, while the fraction of withdrawals in competitions for associate professorships does not decline. The candidates that are now attracted are outsiders, which are less discouraged by the fact of being unknown to the department opening the vacancy: other things being equal, the random selection of committees increases the probability to be connected to one member for any candidate affiliated to an Italian department. Other observable characteristics of candidates (gender, age) do not seem affected by the reform, but the scientific standing of those selected for qualification is on average higher than the pool of applicants.

Since half of the qualified are insiders (i.e. they are affiliated to the department opening the vacancy) both before and after the reform, the effect of being an insider on the probability of being selected does not decline with the reform. Rather, the reform does impact on the effect of having a connection to a commissioner from a university other than the one opening the position. While before the reform it was large and statistically significant it dropped a lot after the reform becoming statistically irrelevant in some of the specification we consider.

In a situation in which the number of candidates is larger and they are less likely to be known in advance to the committee, one would have expected an increase of the relevance of the scientific standing as a criterion for selection. Unfortunately, this does not seem to be necessarily the case, except in competitions in the bibliometric research area to full profes-

sorship. In a similar vein, candidates to associate professorship with high scientific standing become less likely to withdraw their applications after the reform, under the expectation of lower probability of predefined outcomes. The study of withdrawal decisions mirrors the process of selection, suggesting that candidates were capable to correctly anticipate the relative probabilities as before as after the reform.

5. Concluding remarks

Our results confirm that the reform intended to reduce nepotism and localism in the Italian universities has been effective in raising the equality of opportunity for candidates. However this

does not necessarily made selections more metric-based, since there is no clear evidence that the quality of selections has improved. The reform occurred at some costs, which are not easy to evaluate. Randomisation makes the selection of bad commissioners as equally likely as selecting good commissioners. Before the reform a high quality department could arrange a prestigious committee, which would have attracted good candidates and would have made credible an announcement that the best candidates would have been qualified. After the reform, such an announcement would have not been held credible. In other words, randomisation implies levelling to the mean quality, which raises the bottom tail of the quality distribution but compresses the top one.

One could improve the quality of recruitment by introducing minimum requirements for being selected, as it has happened with the new national procedure for qualification, where commissioners and candidate should belong to the top half of the corresponding distribution of scientific productivity (number of papers and number of citations). Participation requirements raise the issue of appropriate measurement of productivity: while in bibliometric research fields there is a widely accepted custom of using citations and impact factors as proxies for quality, in non-bibliometric fields such a consensus does not exist, and therefore it is rather arbitrary to define what a minimum threshold of productivity is.

A clear limitation of our entire analysis is modelling selections as entirely based on scientific productivity. We all know that actual hiring takes into account other dimensions, like teaching, fundraising, agreeableness with colleagues and students, availability for administrative tasks. Unfortunately we do not have information on all these dimensions, which would make our exercise more credible. In the future we hope to be able to enrich our dataset with information on the teaching load assigned to candidates, in order to consider alternative combinations of research and teaching as the driving force for actual selections.

Appendix 1

A1. Homonimity:

We exclude from the analysis candidates whose name is in common with at least another member of the Italian academia. We are not able to control for researchers outside universities, because we do not have information on researchers' names in other research institutions (like CNR) or health centres (like hospitals). However, comparing results in table A1, to results in table 4, columns 6 o 8, it is turns out that at least that the pattern of statistical significance for the estimated coefficients does vary even if in some instances the magnitude of the coefficients does vary.

Table A1: The effect of individual characteristics on candidates' success excluding candidates with at least one homonymous scholar in the Italian academia. Competition to associate professorship.

	Non bibliometric	
VARIABLES	Before	After
H index	0.031*	0.015
	(0.017)	(0.011)
Age	-0.016	0.009
	(0.023)	(0.018)
Age²/1000	0.156	-0.100
	(0.263)	(0.200)
Female	-0.084*	-0.021
	(0.046)	(0.035)
Female×No female evaluators	0.062	-0.018
	(0.072)	(0.067)
Insider	0.404***	0.348***
	(0.062)	(0.061)
Commissioner from the same university	0.294***	0.207***
	(0.059)	(0.061)
Constant	0.501	-0.110
	(0.504)	(0.380)
Observations	532	611
R^2	0.179	0.140
Number of competitions	72	43

*Heteroskedastacity robust standard errors in parentheses, clustered at competition level, including competition fixed effect – statistical significance *** p<0.01, ** p<0.05, * p<0.1*

A2. The effect of distance from the university opening the position on the probability of withdrawing – competitions to associate professorship

Table A2 - The effect of individual characteristics on the likelihood of withdrawal (associate professors); considering only professors with a position in the Italian academy. In columns 5, 6, 7, 8 we take into account also the distance between the university where the candidate has a position and the university opening the vacancy (measured in 100Km).

	Bibliometric		Non Bibliometric		Bibliometric		Non Bibliometric	
VARIABLES	Before (1)	After (2)	Before (3)	After (4)	Before (5)	After (6)	Before (7)	After (8)
Index	-0.042*	-0.020	0.033	-0.041	-0.042*	-0.020	0.033	-0.029
	(0.022)	(0.020)	(0.051)	(0.051)	(0.022)	(0.020)	(0.053)	(0.054)
Age	0.031	-0.059***	0.001	-0.045	0.033	-0.058***	0.004	-0.046
	(0.024)	(0.018)	(0.033)	(0.028)	(0.024)	(0.018)	(0.033)	(0.027)
Age^2 /1000	-0.374	0.665***	-0.057	0.523	-0.399	0.653***	-0.076	0.540*
	(-0.261)	(0.197)	(0.352)	(0.314)	(0.260)	(0.192)	(0.354)	(0.307)
Female	0.005	0.048	0.104	0.024	0,007	0.043	0.105	0.033
	(0,036	(0.029)	(0.066)	(0.069)	(0,036)	(0.030)	(0.065)	(0.066)
Candidate distance					0.015***	0.020***	0.012	0.017*
					(0.005)	(0.004)	(0.011)	(0.009)
Femalex No female evaluators	-0.045	0.043	-0.090	0.011	-0.045	0.050	-0.096	0.024
	(0.061)	(0.049)	(0.144)	(0.127)	(0.061)	(0.049)	(0.143)	(0.131)
External above internalxH index	-0.023	-0.038	-0.099	0.017	-0.21	-0.033	-0.097	0.011
	(0.031)	(0.030)	(0.067)	(0.057)	(0.030)	(0.030)	(0.068)	(0.059)
Insider	-0.382***	-0.323***	-0.353***	-0.254***	-0.316***	-0.245***	-0.309***	-0.193**
	(0.029)	(0.027)	(0.068)	(0.048)	(0.038)	(0.031)	(0.079)	(0.056)
Evaluator of the same university	-0.212***	-0.067**	-0.247***	-0.147**	-0.204***	-0.077**	-0.242***	-0.151*
	(0.039)	(0.033)	(0.062)	(0.071)	(0.039)	(0.035)	(0.061)	(0.072)
Constant	0.105	1.838***	0.633	1.381**	-0.029	1.712***	0.536	1.355**
	(0.536)	(0.425)	(0.770)	(0.621)	(0.529)	(0.410)	(0.783)	(0.612)
Observations	1,480	1,939	352	460	1,480	1,939	352	460
	0.130	0.084	0.135	0.067	0.136	0.095	0.137	0.076
Number of competitions	195	117	72	43	195	117	72	43

*Robust standard errors in parentheses, clustered at competition level (including competition fixed effect) – statistical significance *** p<0.01, ** p<0.05, * p<0.1*

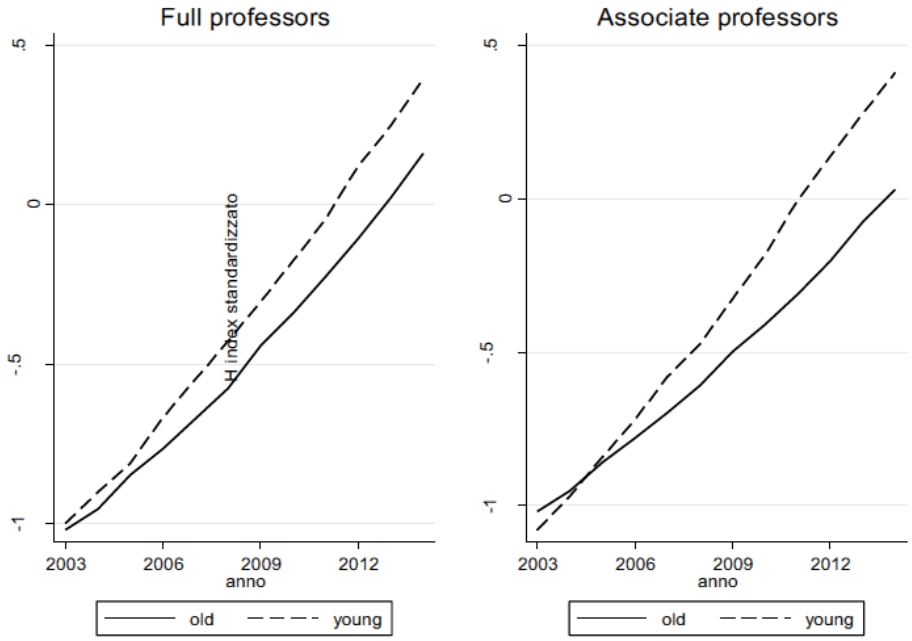
A3. Timing of data collection of H index measures

The analyses we present in this paper make use of an H index for candidates and for commissioners evaluated in 2014. Clearly, this is not the same as the H index of those individuals at the time the competitions took place. To investigate on the possible bias arising from this approximation we have collected data on the H index at the time of the competition for a sub-sample of candidates in bibliometric competitions. This sub-sample is composed by two candidates in each macro research area and in each year of competitions.[20] In figure 1 we observe the trend of H index between 2003 to 2014. As described in section 3, the H index is standardized: we find that the increase of H index is linear even if the trend is slightly different between young and old candidates[21].

[20] There are 14 macro areas of research (the so-called Area CUN), where all Italian academics are assigned to, depending on the research field they are pigeonholed when initially hired. The bibliometric sectors concern only 10 macro areas, which are observed over 5 years covered by our sample (from 2003 to 2006 and 2008): therefore we should have a theoretical sample of 100 candidates, both for full and associate professor competitions. However, we lose 4 candidates for full professorship because of the lack of competitions in one area for two years and we were forced to exclude another area (10 candidates) in competition for associate professorship due to too few candidates.

[21] To define young candidates we observe those applicants who have an age below the median of the age of

Figure 1 - Differences between H index measured in 2014 and H index measured at the time of competition application 96 candidates for full professorship and 90 for associate professorship.



The measurement error is $(h_{ij}-h^*_{ij})$ where h^*_{ij} is the H index at the time of competition and h_{ij} is the H index observed in 2014. We compute the average error for each research area *j* for each year of competition *t*:

$$\mu_{ij} = E[h_{ij}-h^*_{ij}]$$

Then we compute the square difference between the measurement error of the candidate *i* who participated in a competition at the time *t* in the research area *j* and μ_{ij}:

$$(h_{ij}-h^*_{ij}- \mu_{ij})^2$$

and we regress this variable on research area (identified by a dummy variable) and on three dummy for the time of competition (competitions in 2003-2004 or in 2005-2006 or in 2008). We do not find any systematic bias resulting from the use of this approximation.

the sample.

Table A3 - Prediction of the measurement error when using current values of H index instead of past ones

VARIABLES	(1) Associate	(2) Full
Years: 2005-2006	-0.159	-0.154
	(0.138)	(0.177)
Year: 2008	-0.485***	-0.235
	(0.169)	(0.213)
Research areas (*areaCUN*)		
ac2	-0.0343	0.168
	(0.262)	(0.395)
ac3	-0.0569	-0.0257
	(0.262)	(0.338)
ac4	0.184	-0.103
	(0.262)	(0.338)
ac5	-0.127	-0.284
	(0.262)	(0.338)
ac6	0.0597	-0.258
	(0.262)	(0.338)
ac7	0.251	0.777**
	(0.262)	(0.338)
ac9	0.808***	0.104
	(0.262)	(0.338)
ac11	0.0402	-0.0995
	(0.262)	(0.338)
ac8		0.0594
		(0.338)
Constant	0.419**	0.411
	(0.200)	(0.258)
Observations	90	96
R-squared	0.253	0.153

Standard errors in parentheses
**** p<0.01, ** p<0.05, * p<0.1*

A4. Standardizing the H index of applicants and commissioners.

Table A4- -Mean and standard deviation of the H index by research area – applicants to associate professorships – years 2003-2005 (before the reform).

bibliometric	research area (*areaCUN*)	research area	MEAN H index	st.dev. H index	N obs
1	1	Mathematics and informatics	8.2	6.5	239
1	2	Physics	19.6	12.1	159
1	3	Chemistry	19.5	8.2	156
1	4	Earth sciences	10.4	6.2	112
1	5	Biology	18.4	9.3	254
1	6	Medicine	17.1	10.9	763
1	7	Agricultural and veterinary sciences	9.4	6.5	204
1	8	Civil engineering	8.3	4.3	16
1	9	Industrial and information engineering	9.7	5.8	118
0	10	Antiquities, philology, literary studies, art history	2.6	2.3	77
1	11	Psychology	7.1	7.5	41
0	12	Law studies	4.1	7.6	100
0	13	Economics and statistics	9.7	6.0	96
0	14	Political and social sciences	7.7	5.1	84
0	8	Architecture	3.4	6.3	124
0	11	History, philosophy, pedagogy	6.0	5.5	76

Table A5 - Mean and standard deviation of the H index by research area – applicants to full professorships – years 2003-2005 (before the reform).

bibliometric	research area (*areaCUN*)	research area	MEAN H index	st.dev. H index	N obs
1	1	Mathematics and informatics	9.0	5.0	129
1	2	Physics	23.8	15.3	61
1	3	Chemistry	20.0	9.8	111
1	4	Earth sciences	12.8	8.5	32
1	5	Biology	18.3	10.4	155
1	6	Medicine	17.1	12.4	313
1	7	Agricultural and veterinary sciences	8.8	6.1	38
1	8	Civil engineering	10.2	4.4	66
1	9	Industrial and information engineering	13.6	7.0	64
0	10	Antiquities, philology, literary studies, art history	6.6	9.1	40
1	11	Psychology	11.3	8.8	26
0	12	Law studies	7.7	10.2	51
0	13	Economics and statistics	7.8	6.7	10
0	14	Political and social sciences	11.1	10.6	39
0	8	Architecture	8.8	11.7	46
0	11	History, philosophy, pedagogy	7.5	7.5	36

A5. Pure randomization vs election+ randomization. Does it make any difference?

TABLE 6 - The effect of individual characteristics on the likelihood of success (full professors).

VARIABLES	Bibliometric			Non bibliometric		
	Before	After, election+random	After, pure randomization	Before	After, election+random	After, pure randomization
H index	0.033**	0.073***	0.072***	0.024	-0.035	-0.002
	(0.016)	(0.017)	(0.017)	(0.031)	(0.040)	(0.013)
Insider	0.383***	0.366***	0.314***	0.520***	0.485***	0.289***
	(0.037)	(0.057)	(0.046)	(0.076)	(0.113)	(0.048)
commissioner from the same university	0.147***	0.099*	0.036	0.189*	-0.081	0.151*
	(0.031)	(0.052)	(0.038)	(0.103)	(0.062)	(0.084)
Observations	995	541	639	222	85	215
R-squared	0.164	0.208	0.159	0.236	0.215	0.131
Number of ID_competition	100	38	32	38	6	16

*Heteroskedasticity robust standard errors in parentheses, clustered at competition level, including competition fixed effect – statistical significance *** p<0.01, ** p<0.05, * p<0.1*

TABLE 7 - The effect of individual characteristics on the likelihood of success (associate professors).

VARIABLES	Bibliometric			Non bibliometric		
	Before	After, election+random	After, pure randomization	Before	After, election+random	After, pure randomization
H index	0.047***	0.036***	0.047***	0.040**	-0.012	0.051**
	(0.010)	(0.011)	(0.015)	(0.017)	(0.011)	(0.020)
Insider	0.346***	0.289***	0.161***	0.395***	0.409***	0.281***
	(0.026)	(0.030)	(0.029)	(0.059)	(0.090)	(0.069)
commissioner from the same university	0.192***	0.018	0.071*	0.292***	0.173*	0.229***
	(0.030)	(0.024)	(0.042)	(0.055)	(0.093)	(0.065)
Observations	2,062	1,188	1,156	557	347	289
R-squared	0.144	0.132	0.072	0.174	0.141	0.151
Number of competitions	201	64	53	72	24	19

*Heteroskedasticity robust standard errors in parentheses, clustered at competition level, including competition fixed effect – statistical significance *** p<0.01, ** p<0.05, * p<0.1*

A6. Representativeness of our sample

TABLE 8 – Number of competitions by research area in our sample (Padua, Milan – 'Statale', Naples – 'Federico II' and Rome – 'La Sapienza') and in the whole Italian academia - associate professor competitions.

bibliometric	research area (*areaCUN*)	research area	Sample		all universities	
			N	%	N	%
1	1	Mathematics and informatics	31	7.1	174	4.6
1	2	Physics	22	5.1	137	3.6
1	3	Chemistry	32	7.4	150	3.9
1	4	Earth sciences	18	4.2	58	1.5
1	5	Biology	42	9.7	254	6.7
1	6	Medicine	102	23.5	625	16.4
1	7	Agricultural and veterinary sciences	36	8.3	201	5.3
1	8	Civil engineering	4	0.9	98	2.6
1	9	Industrial and information engineering	24	5.5	331	8.7
0	10	Antiquities, philology, literary studies, art history	25	5.8	339	8.9
1	11	Psychology	8	1.8	88	2.3
0	12	Law studies	26	6	400	10.5
0	13	Economics and statistics	17	3.9	407	10.7
0	14	Political and social sciences	14	3.2	125	3.3
0	8	Architecture	12	2.8	147	3.9
0	11	History, philosophy, pedagogy	21	4.8	281	7.4
		Total	434		3815	

TABLE 9 – Number of competitions by research area in our sample (Padua, Milan – 'Statale', Naples – 'Federico II' and Rome – 'La Sapienza') and in the whole Italian academia – full professor competitions.

bibliometric	research area (*areaCUN*)	research area	Sample		all universities	
			N	%	N	%
1	1	Mathematics and informatics	11	4.8	95	3.9
1	2	Physics	6	2.6	75	3.1
1	3	Chemistry	19	8.3	98	4.0
1	4	Earth sciences	6	2.6	44	1.8
1	5	Biology	23	10	152	6.2
1	6	Medicine	58	25.2	365	14.8
1	7	Agricultural and veterinary sciences	19	8.3	114	4.6
1	8	Civil engineering	5	2.2	54	2.2
1	9	Industrial and information engineering	15	6.5	215	8.7
0	10	Antiquities, philology, literary studies, art history	13	5.7	223	9.1
1	11	Psychology	8	3.5	69	2.8
0	12	Law studies	16	7	288	11.7
0	13	Economics and statistics	6	2.6	274	11.1
0	14	Political and social sciences	7	3	80	3.3
0	8	Architecture	9	3.9	90	3.7
0	11	History, philosophy, pedagogy	9	3.9	226	9.2
		Total	230		2462	

TABLE 10 – Number of competitions and number of applications per competition (associate professors).

year	number of competitions			average applicants per competition		
	Italy	4 universities	our sample	Italy	4 universities	our sample
2003	515	72	61	6.9	7.0	8.3
2004	792	68	54	8.5	8.3	8.8
2005	1048	149	135	10.2	10.9	11.3
2006	244	23	23	4.7	4.9	5.0
2007	4	-	-	5.0	-	-
2008	1213	162	161	14.2	18.7	18.7
Total	3816	474	434	10.3	12.3	13.0

TABLE 11 – Number of competitions and number of applications per competition (full professors).

year	number of competitions			average applicants per competition		
	Italy	4 universities	our sample	Italy	4 universities	our sample
2003	330	34	30	6.9	6.7	6.6
2004	530	29	21	8.1	7.2	7.3
2005	667	69	69	10.1	11.0	11.4
2006	193	18	18	3.8	4.6	4.6
2007	4			3.5		
2008	738	90	92	14.6	15.9	16.1
Total	2462	240	230	10.1	11.3	11.7

References

Allesina, Stefano. 2011. Measuring Nepotism through Shared Last Names: The Case of Italian Academia. PLoS ONE, 2011, 6, e21160.

Bagues, Manuel, Mauro Sylos-Labini, Natalia Zinovyeva. 2014. Do gender quotas pass the test? Evidence from academic evaluations in Italy. LEM WP 2014/14

Bagues, Manuel, Mauro Sylos-Labini, Natalia Zinovyeva. 2015. Connections in Scientific Committees and Applicants' Self-Selection: Evidence from a Natural Randomized Experiment. IZA DP n.9594

Checchi, Daniele, Gianni DeFraja e Stefano Verzillo. 2014. Publish or Perish? Incentives and Careers in Italian Academia. CEPR Discussion Paper n. 10084

De Paola, Maria, Michela Ponzo, Vincenzo Scoppa. 2014. Gender Differences in Attitudes Towards Competition: Evidence from the Italian Scientific Qualification. mimeo 2014

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

Durante, Ruben, Giovanna Labartino, and Roberto Perotti. 2011. Academic Dynasties:

Decentralization and Familism in the Italian Academia. NBER Working Paper No. 17572 November 2011.

Godechot, Olivier. 2015. The Chance of Influence: A Natural Experiment on the Role of Social Capital in Faculty Recruitment. Forthcoming in Social Networks

Moss, David. 2012. When Patronage Meets Meritocracy: Or, The Italian Academic Concorso as Cock_fight. European Journal of Sociology, 53, 205-231.

Perotti, Roberto. 2008. L'università truccata, Torino, Italy: Einaudi.

Zinovyeva, Natalia and Manuel Bagues. 2015. The Role of Connections in Academic Promotions. American Economic Journal: Applied Economics, 72(2), 264-292

SENATO DELLA REPUBBLICA

UFFICIO VALUTAZIONE DI IMPATTO
IMPACT ASSESSMENT OFFICE

www.senato.it/ufficiovalutazioneimpatto