WORKING PAPER NO. 507

The Effectiveness of Promotion Incentives for Public Employees: Evidence from Italian Academia

Marco G. Nieddu and Lorenzo Pandolfi

July 2018

University of Naples Federico II University of Salerno Bocconi University, Milan

CSEF - Centre for Studies in Economics and Finance
DEPARTMENT OF ECONOMICS – UNIVERSITY OF NAPLES
80126 NAPLES - ITALY
Tel. and fax +39 081 675372 – e-mail: csef@unina.it
ISSN: 2240-9696

WORKING PAPER NO. 507

The Effectiveness of Promotion Incentives for Public Employees: Evidence from Italian Academia

Marco G. Nieddu[*] and Lorenzo Pandolfi[**]

Abstract
This paper investigates how promotion incentives affect the productivity of high-skilled public employees. In a fuzzy regression discontinuity design, we exploit the three bibliometric thresholds of the 2012 National Scientific Qualification (NSQ), the centralized evaluation procedure awarding the eligibility for career advancements in Italian universities. Specifically, we compare the 2013-2016 research productivity of assistant professors who barely achieve the qualification for associate professor with the productivity of candidates who barely miss it. The former have the incentive to enrich their publication records in order to meet the higher requirements for the full professor qualification by the following round of the NSQ. Conversely, the latter first need to re-apply for the associate professor qualification, thus facing lower promotion thresholds. We find that barely qualified scholars publish significantly more papers – and in journals of comparable quality – than their unsuccessful colleagues. The relationship between the increase in publications and the distance from the expected thresholds for the full professor qualification is inverted-U shaped: promotion incentives are mostly effective when the promotion threshold is neither too difficult nor too easy to meet. Our results emphasize the importance of promotion incentives as an effective tool for public management to enhance the productivity of state personnel. They also provide novel evidence on the responsiveness of scholars to publication-based hiring and promotion schemes.

JEL Classification: I23; J45; M51; O31

Keywords: promotion incentives; public sector; academia; Italy; scientific productivity.

Acknowledgements: We are very grateful to Albrecht Glitz and Gianmarco León for numerous insightful discussions and suggestions. This paper greatly benefited from helpful comments by Niklas Heusch, Tullio Jappelli, Maria Lombardi, Roberto Nisticò, Tommaso Oliviero, Marco Pagano, Michela Ponzo, Marta Santamaria, Annalisa Scognamiglio, Saverio Simonelli, Alberto Zazzaro, seminar participants at CSEF, UPF and University of Cagliari, and participants at the 2018 Petralia Sottana Applied Economic Workshop.

* Università di Cagliari and CRENoS.

** Università di Napoli Federico II and CSEF. E-mail: lorepandolfi@gmail.com

Table of contents

I Introduction

Rewards for good performance are a key tool for firms and organizations to motivate and retain employees. In the private sector, performance-based incentives have been shown to be an effective tool in enhancing productivity.[1] Conversely, the literature on how to motivate civil servants is much smaller and limited mostly to pay-for-performance schemes, with very few studies focusing on promotion incentives. This lack of empirical evidence is concerning, as career-based incentives, in fact, represent the main motivational lever in the hands of public management.[2] Compensation schemes in the public sector are indeed typically rigid and do not easily allow for the inclusion of discretionary performance-based components (Finan et al., 2017). Additionally, in most countries, a significant share of the entire labor force is employed in the public sector, whose productivity therefore represents a key factor for economic growth.[3] Deeper insight into whether and how promotions can effectively incentivize workers in public organizations is thus needed, as the question is important from both a policy and academic perspective.

The aim of this paper is precisely to shed light on this topic. We assess the impact of promotion incentives on the performance of high-skilled public employees by studying whether a quasi-random assignment of different career prospects affects the research productivity of approximately 5,000 assistant professors in Italy. The current Italian public university system represents an ideal laboratory to address our research question, as it is characterized by a clear hierarchical structure and a centralized promotion mechanism based on observable measures of individual scientific production. Moreover, scientific productivity of academics is a topic of particular importance per se. The production of knowledge is recognized to be one of the main engines of economic growth and, in many countries, the main provider of research and education is the state. However, while there is some

[1] The literature on the effectiveness of financial incentives in the private sector includes, among others, Lazear (2000), Gaynor et al. (2004), Shearer (2004) and Friebel et al. (2017). For career-based incentives, see Kwon (2006) and Campbell (2008). See also Lazear and Oyer (2013) for an exhaustive review of related studies.

[2] Also according to Haeck and Verboven (2012), promotions constitute the main source of incentives for public workers. Studying the case of a European university, they find that its labor market is characterized by a strong barrier at the entry level, salaries that evolve independently from external wages and long internal career progressions.

[3] As of 2015, public employment accounts for 18.1% of total employment across OECD countries. In Scandinavian countries, this share increases up to approximately 30%, almost twice that of the US (15.3%) (OECD, 2017)

evidence on the determinants of productivity in the scientific production process, little is known about how academics respond to different recruitment and promotion schemes. We contribute to filling this gap by investigating whether linking career advancements to scholars' publication records – by setting explicit promotion thresholds – can effectively foster research productivity.

In our analysis, we exploit the introduction, in 2012, of a centralized evaluation procedure awarding the eligibility for career advancements – namely, the *National Scientific Qualification* (henceforth NSQ) – on the basis of past performance. We take advantage of a peculiar feature of this procedure: success in the NSQ depends on scholars' past research productivity, measured by three *bibliometric* indicators that are required to be above certain observable and well-defined thresholds. Hence, in a regression discontinuity design with three running variables, we compare the post-2012 research productivity of barely successful and unsuccessful assistant professors applying for the associate professor qualification. While the former can achieve the qualification for a full professorship in the subsequent round of the NSQ, the latter first need to re-apply for the associate professor qualification.[4] Success or failure in the 2012 NSQ thus generates very different promotion incentives. Qualified candidates have the incentive to enrich their publication records in order to meet the higher eligibility requirements for a full professorship in the subsequent round of the NSQ. Conversely, the goal for barely unsuccessful scholars remains meeting the associate professor thresholds, which are, by definition, very close.[5]

Our (triple) regression discontinuity estimates show that achieving the qualification in 2012 – and thus being exposed to higher promotion incentives – has a positive and significant effect on the number of scientific papers published in the subsequent four years. The marginally qualified scholar publishes on average 6 items more than her marginally unsuccessful colleague. This effect is sizable, as it corresponds to a 38% increase with respect to the average number of publications in

[4] Although the system does not explicitly prevent assistant professors from applying directly for the full professor qualification, the probability of succeeding without having already obtained the associate professor qualification is *de facto* very low, as we show in greater detail in Section IV.

[5] By 'unsuccessful scholars', we refer to both the applicants who are denied the qualification and those who withdraw their application before the committee evaluates their applications. In our regression discontinuity framework, the 'barely unsuccessful' candidates are the ones whose *bibliometric* indicators are 'almost' at the threshold. Thus, the gap they must fill in order to overcome the minimum requirements for the associate professor qualification in one of the later rounds of the NSQ is close to zero.

the entire sample. We also find that the average publication quality – proxied by different measures of the journal's prestige – does not exhibit any discontinuity at the multidimensional threshold. The increase in the quantity of publications thus does not occur at the expense of the average quality. Additionally, we provide suggestive evidence that qualified scholars tend to expand the co-author network and receive more citations, compared to their unsuccessful colleagues. When investigating the heterogeneity of our results depending on candidates' gender, we find no evidence of a differential responsiveness to the provision of promotion incentives between male and female assistant professors. However, we find that women who comply with the cutoff rule are less likely to achieve the qualification than their male colleagues, which may suggest the presence of gender discrimination in the evaluation procedure.

Digging deeper into the distributional effects of promotion incentives, we show that the effect of the qualification on the number of publications is heterogeneous, depending on scholars' distance from the (expected) future promotion thresholds. The relation between the increase in productivity and the distance from the thresholds for the full professor qualification is inverted-U shaped: the effect is the strongest for candidates in the middle of the distribution, while it is not statistically significant in the group of candidates for whom the minimum requirements to earn the qualification for a full professorship are either too close or too far. This result is consistent with the view that incentives are mostly effective when "the promotion is possible, but neither too hard to achieve, nor too easy" (Lazear and Gibbs, 2014, p.269). Besides shedding light on the distributional consequences of the promotion incentives, this evidence lends important support to our empirical strategy. Most of the competing explanations for our main results – for instance, a motivational effect arising from succeeding in the NSQ, or other changes in scholars' daily life occurring after the achievement of the qualification – can hardly be reconciled with the heterogeneity of the estimated effects depending of the distance from the future promotion threshold.

Several additional pieces of evidence corroborate the robustness of our results and confirm that promotion incentives are the main driver of our findings. First, we take advantage of the longitudinal dimension of our data to rule out that the observed discontinuity in the post-call research productivity is driven by a decline in performance of discouraged, unsuccessful candidates rather than by an increase in the productivity of qualified scholars. Second, we replicate our analysis with the sample of associate professors applying for the full professor qualification in 2012. Consistent with the fact that promotion incentives vanish once the top ladder of the academic hierarchy is reached, we do not observe any discontinuity in this alternative sample. Further, we show that the increase in productivity of qualified candidates occurred already in 2013, immediately after the achievement of the NSQ and prior to the actual promotion to associate professor. Hence, the timing of the effect suggests that our results are not driven by a variation in teaching duties, research funds, and other aspects of scholars' routine that may change with career advancement. Lastly, we exploit the between-field heterogeneity in the share of candidates qualified for an associate professorship on the total number of associate professors already employed in each sector to show that competition for vacancies among qualified candidates is not driving our findings.

This study contributes mainly to the personnel economics literature and, more precisely, to the stream of studies focusing on the design of incentives in the public sector. As highlighted by Finan et al. (2017), public sector pay schemes are typically flat, with salaries that are mechanically determined by seniority and position and rarely linked to workers' performance.[6] Also because of this, most of the related studies focus on performance-pay programs implemented in the context of randomized control trials and typically do not involve high-skilled workers. For instance, Muralidharan and Sundararaman (2011) evaluate the impact of a randomized performance-pay program in India and find that linking teachers' pay to students' test scores has a positive effect on learning. In another randomized experiment, Duflo et al. (2012) show that performance-related pay lowers absenteeism among Indian teachers, which in turn translates into better students' performance.[7] This literature also highlights a potential pitfall of implementing performance-pay schemes in the public sector: as public jobs typically involve multiple tasks, financial incentives based on the performance in a specific task can reduce workers' effort in another (Baicker and Jacobson, 2007; Glewwe et al.,

[6] The same argument is also made by Khan et al. (2018) in their study on the efficacy of performance-based job posting schemes for government employees.

[7] Other studies focusing on the role of financial incentives in the public sector are those by Lavy (2002), Gertler and Vermeersch (2013), Dal Bó et al. (2013), Olken et al. (2014), and Khan et al. (2016). See Finan et al. (2017) for an exhaustive review.

2010).

Therefore, while some studies have examined the effectiveness of monetary incentives on the productivity of public employees, the existing literature has overlooked the role of promotion incentives. The latter constitute a complicated subject of study – even in the private sector – since promotion incentives can hardly be implemented in the context of randomized controlled trials, and quasi-experimental evidence on the topic is rare. To our knowledge, the only study that explicitly focuses on the relationship between promotion incentives and workers' productivity in the public sector is that by Karachiwalla and Park (2017).[8] The authors exploit the Chinese system regulating teachers' career advancement to test the prediction of a tournament model of promotions and show that promotion incentives are associated with higher levels of performance. Related to this is the study by Checchi et al. (2014), who develop a model of career concerns in academia whose predictions are consistent with data on the publications of Italian scholars from 1990 to 2011. Unlike these studies, we exploit quasi-experimental variation in promotion incentives – coming from explicit promotion thresholds implemented within an actual governmental policy – to uncover the causal link between promotion incentives and the productivity of public employees.

More broadly, given our focus on academia, this study also speaks to the literature focusing on the research productivity of scholars. Some studies in this stream of literature investigate the relative importance of human and physical capital in the scientific production process (Waldinger, 2016), the relevance of peer effects in science (Waldinger, 2012; Borjas and Doran, 2015b), and the importance of having access and being exposed to the scientific research frontier (Iaria et al., 2018). Some others focus on the effect of achieving tenure (Faria and McAdam, 2014) or prestigious awards (Borjas and Doran, 2015a) on research outputs. We contribute to this literature by uncovering the responsiveness of scholars to hiring and promotion schemes explicitly based on past research performances.

Finally, our paper is also related to the recent literature focusing on the centralized evaluation

[8] Additionally, Ashraf et al. (2018) study the role of career prospects on the recruitment of workers in the public sector (health) in Zambia and on the quality of the service delivered. However, that study focuses on the selection channel.

systems that have been introduced in the last decade in several European countries to regulate access to public university positions. Similar to the Italian NSQ are, among others, the *Acreditación* in Spain and the *Habilitation à diriger des recherches* in France. All the related studies focus on the functioning of the evaluation process and, more specifically, on the role of gender (Bagues et al., 2017; De Paola and Scoppa, 2015; De Paola et al., 2017) or of direct connections between evaluators and candidates (Zinovyeva and Bagues, 2015). None of them examines the potential implications for scholars' productivity, as we do in this study.

The rest of the paper is organized as follows. Section II describes the regulatory framework and the key features of the NSQ. The data used for the empirical analysis and the identification strategy are reported in Sections III and IV, respectively. We then present the first-stage estimates in Section V. Section VI contains our main results, together with some additional robustness tests. In Section VII, we dig deeper into the promotion incentives mechanism by exploring the the distributional effects of promotion incentives and discarding potential competing hypotheses for our findings. Finally, Section VIII concludes.

II Institutional Setting

In 2010, the Italian Ministry of Education, University and Research (MIUR) deeply reorganized the public university system through the so-called Gelmini reform. The latter introduced a new recruitment and promotion system regulating the access to the two top ranks of the academic hierarchy: the associate professorship and the full professorship.[9] Until that time, the hiring and the promotion processes were fully decentralized and each academic department had complete discretion over the selection procedure. Since the reform came into force, however, earning an associate or full professorship is conditional on having achieved a qualification – the National Scientific Qualification (*Abilitazione Scientifica Nazionale*) – which is awarded by national committees in a centralized

[9] The hierarchical structure of Italian universities consists of three main ranks: assistant, associate and full professors. Until 2010, the three positions were all tenured and assistant professors were hired under permanent contracts. After the reform, instead, assistant professors are hired under fixed-term contracts.

evaluation process. The first round of the NSQ took place in 2012 and was followed by two more rounds, in 2013 and 2016-2018.[10]

By achieving the NSQ, scholars gain the mere eligibility for associate and full professorships, while actual hirings still occur at the university department level in a decentralized stage. Coherent with the rationale of the reform, which is to promote research activity and limit local favoritism, the introduction of the NSQ restricts the access to local competitions to candidates whose academic *curriculum vitae* satisfies minimum standards established at the national level. Applicants are evaluated by a committee of five scholars randomly drawn from a list of eligible full professors affiliated with Italian and non-Italian universities. Both the evaluation criteria and the committee composition vary depending on whether a candidate is applying to the NSQ for the associate or the full professorship and on her research field. Academic fields are mapped into 184 different competition sectors, further grouped by the Ministry in two broad macro areas: *bibliometric* (including all disciplines of mathematics, physics, chemistry, earth sciences, health sciences, agronomy and veterinary, engineering and architecture, and psychology) and *non-bibliometric* sectors (humanities, economics, political sciences, and law).

The committee is in charge of screening the items of each candidate's *curriculum vitae* in order to decide whether to award the NSQ or not. The main aspect that committees must take into account when evaluating a candidate is her publication record, measured by three observable and well-defined productivity indicators. In *bibliometric* sectors, these indicators are i) the number of articles published in scientific journals, ii) the number of citations received, and iii) the scholar's h-index. In *non-bibliometric* sectors, they are i) the number of monographs, ii) the number of book chapters and articles in scientific journals, and iii) the number of articles published in a list of A-ranked journals. All indicators are calculated over the ten years prior to the NSQ call and normalized by a candidate's academic age. Other criteria include the participation in national or international projects, editorial activities, fellowships, and awards. Although the Ministry allows committees to decide autonomously the weight assigned to each of the aforementioned elements, it

[10]Differently from the previous two, the 2016-2018 round consists of multiple calls which were opened every four months over a two-years time window.

also explicitly states that the three productivity indicators should constitute the key criteria.[11]

In particular, the ministry defines specific minimum thresholds for attaining the qualification in each field. These standards are set by looking at the publication records of associate and full professors already employed in the Italian university system. In order to achieve the associate (full) professor qualification in a bibliometric sector, a candidate must score above the median associate (full) professor in her sector **in at least two out of the three** productivity indicators. A similar one-out-of-three rule holds for the non-bibliometric sectors. These rules represent a (almost) necessary but not sufficient condition to achieve the qualification since committee members might deliver a negative judgement even when all of a scholar's indicators surpass the relevant thresholds. Moreover, they also have the right to deviate from the aforementioned rule by awarding the qualification to candidates who do not comply with the productivity requirements. Nonetheless, this latter possibility is allowed only in case of an extremely positive evaluation of the other elements of the *curriculum*.[12]

The first round of the NSQ opened between June and July 2012, when both the call for commissioners and that for candidates were published. In August, the ministry released the sector-specific cutoff values for the each of the three productivity indicators. The deadline for candidates to apply was set for the 20th of November. After this date, the ministry made public candidates' scores, calculated by the ministry, and the list of commissioners in each field, randomly drawn from the list of eligible full professors. Candidates had the right to withdraw their applications until February 2013. This option was particularly important since a negative assessment by the committee in the 2012 round of the NSQ implied that a candidate could not apply to the subsequent one.[13] Thus,

[11]Each committee is composed of four full professors at Italian universities and one at a university in a different OECD country. The eligibility requirements for commissioners are similar to those for candidates: when considering the aforementioned productivity indicators, only full professors above the median in all of the three indicators can be part of the evaluating committee for a given field.

[12]Figure A2 in the Appendix depicts the extent of the deviation from the two-out-of-three rule across different academic fields. The green bars describe the proportion of candidates who obtained the qualification in the 2012 NSQ despite not complying with the two-out-of-three rule. On average, fewer than 15% of candidates who did not comply with the two-out-of-three rule achieved the qualification in 2012.

[13]Candidates who received a negative judgement by the committee were prevented from applying for the qualification for the same rank and in the same competition sector for two years, which implies that rejected candidates could not apply to the following one or two rounds, depending on the exact timing of the calls.

applicants could decide whether to undergo the evaluation or not after having observed their scores, the cutoff values, and the composition of the evaluation committee. Most committees completed their work and published the outcome by June 2013, while in few cases, the evaluation process took until December 2013.[14]

III Data and Sample Description

In this study, we combine different administrative and publicly available data sources to build a unique and comprehensive dataset containing, for each candidate for the 2012 NSQ, i) the score in each of the three productivity indicators and the outcome of the qualification procedure; ii) the academic position and affiliation at the time of the call; iii) the complete publication record from 2007 to 2016.

The list of applicants to the 2012 round of the NSQ is obtained from the MIUR website. The administrative records include information about each candidate's application(s), that is, the competition sector, the scores in the productivity indicators, the sector-specific cutoffs, and the final outcome of the evaluation procedure. We merge these data with the professor census, which covers all assistant, associate, and full professors employed in the Italian public university system in 2012-2016.[15] This longitudinal database allows us to determine, for each applicant, her position, department of affiliation, academic field as of 2012, and later promotion patterns. Since the NSQ system allows for multiple applications per candidate, in our baseline specification, we consider each candidate's 'best' application in terms of distance from the cutoffs. However, as shown in Section VI.b, results are robust to considering, for each scholar, the outcome of her application in the competition sector in which she was already employed at the time of the call.

Our measures of research productivity come from the *Scopus* database, the largest repository of

[14] Importantly, as also discussed in Section III, our dataset covers all participants to the 2012 call at the time of the deadline (November 2012), thus also including withdrawn applications.

[15] A tiny share (4% in our final sample) of professors in the census are employed in a few private Italian universities which however are subject to the same regulatory framework that applies to public institutions as regards hiring and promotions.

peer-reviewed literature. We query the *Scopus* archives in order to retrieve each scholar's complete publication record. For each item, we obtain the cover type (article, conference paper, book chapter or review), author's affiliation, publication date, journal name, and the full list of coauthors. Then, we use this information to build a panel dataset at the scholar-by-year level, including measures of the quantity of publications, the quality of the journals in which they are published, and the citations received. The main journal-specific quality indicator is the 2012 *CiteScore index*, which provides a weighted average of the citations received by each journal in a given year. In order to account for the wide heterogeneity between the different academic fields, we look both at the overall *CiteScore index* and at its within-field counterpart, the *CiteScore journal percentile*. Furthermore, we exploit two alternative measures of journals' prestige: the *SJR* (Scimago Journal Rank) and the SNIP (Source Normalized Impact per Paper).[16] The citations received by each published paper are counted as of July 2017.

Overall, approximately 40 thousand researchers took part in the first round of the National Scientific qualification. We discard applicants in non-bibliometric sectors where the cutoff rule is not strictly enforced by most of the committees.[17] In subjects such as humanities, law, political sciences, and economics, where the number of publications per year is typically lower, the thresholds are often very close – or even equal – to zero.[18] Therefore, since more than 90% of the applicants in non-bibliometric sectors satisfy the corresponding one-out-of-three rule, compliance with the latter constitutes a very poor proxy for candidates' quality. Moreover, the resulting lack of observations below the cutoff(s) would not allow us to implement our regression discontinuity design in such fields.

Out of the 20 thousands candidates in bibliometrics sectors – including mathematics, physics,

[16]More precisely, the 2012 *CiteScore* index is computed as the total number of citations received in 2012 by documents published in the three years before, divided by the total number of documents published over the same period. The *CiteScore journal percentile* ranks the journals belonging to each field according to their *CiteScore* index. The *SJR* and *SNIP* indicators are computed in similar way to the *CiteScore index*, thus making them a weighted average of the citations received in a given year by documents published in the three previous years. However, weighting procedures differ from those used to construct the *CiteScore* index.

[17]For 3000 out of the total number of participants in the NSQ it was not possible to identify a unique best application and, consequently, a unique sector. We also exclude them from our analysis.

[18]Given the way cutoffs are established, a threshold equal to zero means that the score of the median associate (or full) professor in that competition sector, for that specific indicator, is equal to zero.

chemistry, earth sciences, health sciences, agronomy and veterinary, engineering and architecture, and psychology – about two thirds applied for the associate professor qualification.[19] Given our focus on promotion incentives, we limit the analysis to the subset of candidates who are already employed as tenured assistant professors at the time of the deadline. This group (7000 scholars) accounts for about 45% of the total number of applicants to the associate professor qualification. The remaining share of applicants consists of researchers working for non-university institutions in Italy or abroad, academics affiliated with non-Italian universities, and young untenured scholars.[20] The wide coverage of *Scopus* allows us to detect a unique author identifier for 97% of the candidates in this sample. For the remaining 3% of the scholars, it could be either the case that none of their publications are recorded in the database or that homonymies and misspelled names result in an unsuccessful merge. Lastly, in our baseline specification we discard within-sector outliers, and observations belonging to competition sectors with fewer than 30 applicants thus ending up with a final sample of 4920 scholars.[21]

A detailed description of the sample is presented in Table I. A significant proportion of candidates for the associate professor qualification were relatively experienced: the average academic age – that is, the number of years since the first publication appears in *Scopus* – is approximately 16 years, as of 2012. Moreover, they published on average 12.40 papers in the four years prior to the NSQ. Slightly less than 60% of candidates achieved the NSQ, whereas two-thirds of them satisfied the two-out-of-three rule. Lastly, in the group of academic fields in our sample the number of collaborations is relatively high both before and after the NSQ: only 2% of the papers published by the assistant professors in our sample is single-authored, and the average number of coauthors per publication is 10.82 (8.27) in the four years following (prior to) the NSQ.

[19] In Section VII we also present our main equation estimated on the sample of associate professors who participated in the NSQ for full professorship in 2012. The process for selecting this sample of applicants to the full professor qualification (4,866) follows the one for candidates to the associate professor qualification.

[20] We exclude this group of applicants because we have no information about their employment status in 2012. Thus, we are unable to distinguish between scholars who are employed in Italian universities, although not tenured, from those who work in other institutions and for whom achieving the NSQ does not imply a variation in promotion incentives.

[21] In Section IV, we explain how we determine outliers and why we eliminate them; additionally, in Section VI.b, we test the robustness of our results to adopting different sample restrictions.

Figure I plots candidates according to the compliance with the two-out-of-three rule – those in the upper-right quadrant, as the figure is drawn for the case when the h-index is below the cutoff – and according to the final outcome of the qualification procedure. The limited degree of fuzziness in Figure I.b confirms that the two-out-of-three rule constitutes a determinant criterion for awarding the qualification in bibliometric fields. Moreover, they also show how the mass of observations concentrates around the multidimensional cutoff, particularly around the intersection of the zero-distance axes. This finding is not surprising since the threshold values are computed by looking at the median associate professor in each field. As discussed in detail in the following section, this particular feature of our data implies that, although local, the effect is estimated in the neighborhood of the representative scholar in each field.

IV Empirical Strategy

We exploit the cutoff rule implemented within the NSQ to determine whether a quasi-experimental provision of promotion incentives significantly affects the productivity of a large sample of high-skilled public employees, such as academics. More precisely, in a regression discontinuity framework, we compare the post-call research productivity of barely successful and unsuccessful assistant professors who participated in the 2012 NSQ call. Before describing the details of the empirical methodology, we discuss by what means achieving or missing the qualification exposes candidates to different promotion incentives in the form of different promotion thresholds to be met in the future rounds of the NSQ.

IV.a Promotion incentives in the NSQ

The NSQ introduces explicit thresholds that scholars have to meet in order to gain the eligibility for career advancements. Such multiple-step procedure with clear and well-known criteria for promotions therefore entails significant dynamic incentives by making possible, for those ranked lowest, to climb the academic ladder in a few years.[22]

Although earning the qualification does not immediately imply an advancement to the next rung of the career ladder (associate professorship), it sets a new attainable goal to be achieved in the subsequent call of the NSQ: the qualification for full professorship. In order to fulfill this goal, qualified candidates need to enrich their publication record to meet the corresponding bibliometric requirements, that is, the number of articles and citations and the h-index required to achieve the full professor qualification. Figure II shows that approximately one-third of the assistant professors who successfully took part in the NSQ for associate professor in 2012 were indeed able to also earn the eligibility for a full professorship by 2016.[23]

Barely unsuccessful scholars, however, are not exposed to the same high-powered incentives. Although not explicitly ruled out by the institutional setting, in fact, it is extremely unlikely for candidates who do not hold the associate professor NSQ to earn the full professor NSQ. The share of assistant professors who succeeded in the qualification for a full professorship without earning the intermediate step in 2012 is indeed very small (4.5%). Moreover, this number drops to 1% for the 2013 and 2016 rounds. Thus, failure in 2012 implies that a candidate will have to re-apply for the qualification for an associate professorship in the future and, for the marginal unsuccessful applicant, this goal does not require a substantial effort increase since her productivity scores are already very close to the relevant thresholds. Promotion incentives for this subset of candidates are clearly weaker or even absent.[24]

In principle, those scholars who have to postpone their career progression could be forward

[22] Although the NSQ might not have been explicitly designed to introduce career-based incentives, as highlighted by Lazear and Gibbs (2014), promotions may constitute an "accidental incentive system“ (p. 262). The perspective of career advancement within the organization could enhance employees' motivation and is inextricably linked to the existence of a hierarchical structure.

[23] This share does not necessarily match the one of assistant professors who obtained a chair shortly after 2016, as actual promotion also depends on universities' turnover and budget constraints. However, the decentralized stage is characterized by limited competition – promotions happen mostly within the initial department of affiliation – and qualified candidates move up the academic ranking smoothly: more of two-thirds of successful applicants in 2012 were actually promoted by the end of 2016. We further discuss the importance of the decentralized stage in our setting in Section VII

[24] Candidates who withdraw their applications after observing their scores and relevant cutoffs can re-apply in the first subsequent round of the NSQ. Rejected candidates, instead, must wait at least two years to apply again for the associate professor qualification in the same competition sector.

looking, already targeting the requirements for the full professor qualification. However, they face more uncertainty, as they have to rely upon the stability of the institutional setting over a longer horizon. Additionally, since the ministry sets the minimum thresholds using the median scholar as the reference point, a new inflow of full professors is likely to affect the productivity distribution and, consequently, the future realization of the cutoffs. These two sources of uncertainty weaken the incentives to target the eligibility requirements for the full professor qualification before succeeding in the associate professor one.

IV.b A triple (fuzzy) regression discontinuity design

Our regression discontinuity strategy exploits the discontinuous jump in the probability of obtaining the qualification, arising when two of the three indicators crosses its relevant threshold. By fully modeling the two-out-of-three rule with three forcing variables – the productivity indicators – we are able to define a three-dimensional cutoff, that is, a hyperplane that is the $\mathbb{R}^3$ equivalent of the standard single-variable frontier. Therefore, one should picture the discontinuity in the probability of receiving the NSQ around the 3-D frontier as a pooled or combined version of the smaller, single-variable, discontinuities. Since the compliance with the two-out-of-three rule alone does not represent a sufficient condition to achieve the qualification, the probability of receiving the treatment will jump by less than 100% when crossing the multidimensional cutoff. Hence, our empirical strategy relies on a (triple) fuzzy regression discontinuity design with multiple sector-specific cutoffs.

Formally, let us define the assignment variables – number of articles, number of citations and h-index – as x_{i1}, x_{i2} and x_{i3}, respectively. Then, G_{iks} is an indicator function that equals one when score k of candidate i belonging to competition sector s is strictly above the cutoff m, that is

$$G_{iks} = \begin{cases} 0 \text{ if } x_{iks} \leq m_{ks} \\ 1 \text{ if } x_{iks} > m_{ks} \end{cases} \quad \text{for each } k \in \{1, 2, 3\}.$$

The indicator D_{is} thus describes the aforementioned two-out-of-three rule:

$$D_{is} = \begin{cases} 0 \text{ if } \sum_{k=1}^{3} G_{iks} < 2 \\ 1 \text{ if } \sum_{k=1}^{3} G_{iks} \geq 2. \end{cases}$$

Consequently, our first-stage equation is

$$Q_{is} = \alpha_0 + \alpha_1 D_{is} + f(x_{iks} - m_{ks}) + \alpha_2 Z_s + \nu_{iks}, \tag{1}$$

where Q_{is} is an indicator that equals one when a candidate achieves the qualification, $f(x_{iks} - m_{ks})$ is a flexible nonlinear function of the distance of the running variables from the threshold(s) (including 2[nd] order polynomials of the three variables and all possible interactions), and Z_s are sector-specific fixed effects. Analogously to a 'canonical' RD design – with a single running variable and single cutoff – the coefficient α_1 measures the discontinuous jump in the probability of achieving the qualification that arises when a candidate complies with the cutoff rule. More precisely, α_1 captures a weighted average of the discontinuity in the probability of achieving the qualification when crossing the frontier hyperplane from all the octants not satisfying the two-out-of-three rule.

This discontinuity in the probability of obtaining the qualification is then used as an instrumental variable to estimate our second-stage equation, which is

$$Y_{is} = \gamma_0 + \gamma_1 \hat{Q}_{is} + f(x_{iks} - m_{ks}) + \gamma_2 Z_s + \eta_{iks}, \tag{2}$$

where γ_1 is the local average treatment effect (LATE) of achieving the NSQ in 2012 on any of our measures of scientific production Y_{is}, computed in the post-call period. The corresponding reduced form equation is

$$Y_{is} = \beta_0 + \beta_1 D_{is} + f(x_{iks} - m_{ks}) + \beta_2 Z_s + \eta_{iks}, \tag{3}$$

where β_1 measures the intention-to-treat (ITT) effect of complying with the two-out-of-three rule. The interpretation of γ_1 and β_1 in this multidimensional regression discontinuity framework is analogous to that provided for the α_1 coefficient of the first stage: they capture a weighted average of the effect of crossing the 3-D frontier from all the neighboring octants.

To account for the heterogeneity between different academic fields, we allow $f(\cdot)$ to be fully flexible across sectors in Equations 1, 2 and 3 by interacting each assignment variable – centered around its sector-specific cutoffs –, their squared values, and their first and second degree interactions, with the competition sector dummies.[25] Because of both the complexity of the framework and the lack of a standard procedure to compute joint bandwidths in a multidimensional regression discontinuity design with multiple cutoffs, our preferred specification is a fully-parametric one. To reduce the weight of potential outliers, we exclude candidates in the top decile or in the bottom percentile of the distribution of the distance from the cutoff. Then, in Section VI.b, we show that results are robust to considering a linear specification within an arbitrary range around the zero-distance cutoff(s) and to adopting alternative sample restrictions to deal with outliers.

Finally, it is important to stress that our identification strategy is less vulnerable to the main criticism usually made for regression discontinuity designs, namely, the locality of the estimated effect. The estimated discontinuity is indeed a weighted average of the discontinuities along the three different frontiers, one for each productivity indicator. Furthermore, cutoff values are set by looking at the median professor in each competition sector. As a result, a large mass of observations is concentrated around the three-dimensional frontier – as highlighted in Figure I – and the marginal candidate in this setting is a representative scholar in her field.

[25] The whole evaluation process should be seen indeed as a combination of many small selections, since candidates face different thresholds and committees depending on the competition sector for which they applied. Figure A1 in the Appendix shows the extent of such across-field heterogeneity in the cutoff values: in many subfields in medicine and physics the median number of articles among associate professors, over the 2002-2012 period, is above 40 articles, while it is often below 10 in mathematics and engineering.

IV.c Validity of the RD design

Our identification strategy relies on two main assumptions: 1) the probability of achieving the qualification jumps discontinuously at the multidimensional cutoff describing the two-out-of-three rule; and 2) the joint distribution of the running variables does not exhibit any bump immediately in the neighborhood of the same cutoff. Furthermore, in a full-parametric, multidimensional regression discontinuity design, special attention should be devoted to possible misspecification issues (3). While the satisfaction of Assumption 1 is discussed in Section V, we address 2 and 3 here.

Testing the validity of Assumption 2 is crucial to discarding two potential threats for our identification strategy: manipulation and sample selection. Regarding the former, the possibility for candidates to manipulate their publication records in order to meet the minimum standards seems remote since both individual scores and thresholds are computed by the Ministry of Education, University and Research. The Ministry collects candidates' full publication records from their application webpage and cross-validates each research item by querying the two largest databases of peer-reviewed literature: *Scopus* and *Web of Science*.[26] Moreover, because of the short time frame between the publication of the call and the application deadline, it is unlikely that scholars would have the time to adjust their publication records to meet the established requirements. Regarding selection, a positive jump in the density could also reveal that scholars who decide to participate in the NSQ without complying with the two-out-three-rule constitute a selected sample. For instance, one potential concern could be that scholars below the cutoff were disincentivized from applying given that a negative evaluation by the committee would have prevented them from participating in the subsequent round of the NSQ. However, it is important to remark that candidates were given the opportunity to withdraw their application after having observed their precise scores and the composition of the committee, and prior to the evaluation itself (that is, by February 2013). Hence, applying to the 2012 NSQ was relatively costless, even for those below the thresholds, and selection concerns should be limited as our sample of candidates is based on the list of applications at the time of the deadline (November 2012) and thus includes withdrawn applications.

[26]More precisely, a ministerial agency (ANVUR) computes both the individual scores and the thresholds.

We formally test whether the distribution of candidates is discontinuous around the cutoff. In Figure III, we report the frequencies, as well as the density and confidence intervals computed following McCrary (2008), for each of the three forcing variables centered around the cutoff. None of the running variables exhibits a significant (at the 10% level) discontinuous jump in the density in the neighborhood of the zero-distance from the cutoff. Point estimates (standard errors) of the density test are 0.085 (0.077) for the number of articles, 0.083 (0.075) for the number of citations and 0.043 (0.046) for the h-index. Furthermore, Figure A3 in the Appendix provides additional support for the assumption that scholars do not endogenously sort or select around the threshold. The robust bias-corrected manipulation test proposed in Cattaneo et al. (2017) delivers non-significant estimates for each of the three running variables (number of articles: T=-0.79, p-val=0.43; number of citations T=-0.27, p-val=0.78; h-index: T=-0.96, p-val.=0.33).[27]

Finally, our fuzzy regression discontinuity design also relies on the assumption that scoring above the median professor in two out of the three bibliometric indicators should have no impact on future scientific productivity other than that passing through the achievement of the NSQ. It seems, however, extremely unrealistic that other confounding factors or policies could drive the observed jumps at such particular cutoffs.

V First Stage

A crucial condition must hold to implement our empirical strategy: overcoming the bibliometric thresholds and satisfying the two-out-of-three rule must result in a discrete jump in candidates' probability of achieving the qualification. In this section, we show that this is indeed the case. In Table III, we report both the estimates of the first-stage equation when considering each of the three bibliometric indicators and the corresponding cutoffs, separately (Columns 1 to 6), and when exploiting the three running variables simultaneously, as formalized by Equation 1 (Column 7). The estimated coefficient from this triple-RDD – that is, our preferred specification – shows that

[27] All estimates are obtained using the Stata package described in Cattaneo et al. (2018).

compliance with the bibliometric two-out-of-three rule discontinuously increases the probability of achieving the qualification for an associate professorship by approximately 30 percentage points. The magnitude of the first stage confirms that commissioners attribute a strong weight to the compliance with the two-out-of-three rule when making their decisions.[28]

The single-RDD estimates are also positive and significant in all specifications, consistent with the graphical evidence in Figure IV. In Columns (1), (3), and (6), we estimate the discontinuity in the probability of achieving the qualification when passing each of the three bibliometric threshold – the number of articles, the citations, and the h-index – , assuming a quadratic functional form on the entire support and including both academic field fixed effects and field-specific interactions. In Columns (2), (4), and (6), we replicate the same estimates assuming a linear functional form within the MSE-optimal bandwidths. In this case, to take into account the wide between-field heterogeneity in candidates' average productivity, we use as running variables the relative distances from each threshold, that is, the original running variable divided by the threshold itself. By doing so, we are also able to compute three optimal bandwidths, expressed in relative terms, which can be used across the different fields.[29] The estimation results are very close to their fully parametric counterparts. Of course, the magnitude of each single-RDD coefficient is lower than that resulting from the triple-RDD estimation since the former measures the discontinuous jumps around each single threshold regardless of whether the specific indicator is pivotal for the compliance with the two-out-of-three rule. Hence, estimating three standard, single-forcing variables RDD would not account for the compliance (or defiance) with the other two requirements, thus increasing the degree of fuzziness. This is precisely the reason why we adopt a triple-RD design, in which the α_1 coefficient of Equation 1 should be interpreted as a combined version of three smaller discontinuities.

[28]The corresponding estimates for the sample of candidates to the NSQ for full professorship are presented in Table A1 in the Appendix.

[29]Specifically, the optimal-MSE bandwidths are computed following Calonico et al. (2014) for each of the three relative distances, separately.

VI Results

VI.a Quantity of publications

Table IV reports the main result of our empirical analysis: achieving the qualification for an associate professorship in 2012 – thus being provided with higher promotion incentives – has a positive impact on the number of papers published in the subsequent years. The local average treatment effect (LATE) of achieving the qualification on the number of scientific publications over the 2013-2016 period corresponds to 6.5 publications and is 3.25 times larger than the intention-to-treat (ITT) effect of complying with the two-out-of-three rule (which is equal to 2 publications). Both the LATE and the ITT coefficients are statistically significant at the 1% level. The estimated effect is sizable and corresponds to approximately 40% (LATE) of the sample average number of publications over the same period. By looking at the different publication types, we find the effect to be driven mostly by an increment in the number of published articles and, to a smaller extent, reviews and conference papers.

In principle, the estimated effect could be due not only to the increased productivity of barely qualified scholars but also to a decline in publications by narrowly unsuccessful candidates. This latter group of scholars might indeed become frustrated and discouraged or could revise their research production function after missing the qualification. In order to disentangle these two hypotheses – the discontinuity being driven by marginal successful or unsuccessful applicants – we exploit the panel dimension of our dataset and replicate our baseline estimation using the yearly number of publications before and after the first call of the NSQ as the dependent variable of interest.

Figure V reports the estimated LATE of the qualification on the number of publications for each year between 2007 and 2016. Blue diamonds describe the evolution of the number of publications for candidates who marginally missed the qualification in 2012, while red circles depict the corresponding trend for those who barely met it. The vertical distances between the two trends represent the estimated discontinuity in each year (the estimated coefficients are reported in Table A4 in the Appendix). The annual productivity of narrowly unsuccessful scholars remains constant in the post-NSQ period, while that of barely successful applicants exhibits a significant rise. This effect is persistent, large in magnitude, and significant for the whole post-call period, with the only exception of 2015 when the discontinuity is still positive but the larger variance in the data lowers its significance.

The estimates in Figure V and Table A4 also lend strong support to our identification strategy. For the entire pre-NSQ period, the difference between treated and non-treated individuals is close to a precise zero. Hence, the results are not driven by any *ex ante* difference between candidates or by a possible misspecification of the functional form assumed when estimating the relation between the treatment and outcomes.

Additionally, we investigate whether the effect of passing the NSQ in 2012 is heterogeneous depending on candidates' gender and academic field. The estimates are presented in Table V and A5, respectively. When looking at gender heterogeneity, we find the LATE to be homogeneous across female and male candidates: the promotion incentives associated with the achievement of the qualification are equally effective, regardless of gender. However, we find a negative and significant coefficient for the interaction between the female indicator and the one for compliance with the bibliometric rule when estimating the first-stage equation. Hence, women who satisfy the two-out-of-three rule are less likely to achieve the qualification than men with comparable publication records. This result is consistent with the evidence provided by Bagues et al. (2017) and De Paola and Scoppa (2015) – who also document that female candidates have lower success rates in the Italian qualification procedure – and could be due to gender discrimination.

We find moderate evidence of between-subject heterogeneity. The interaction coefficients of the field-specific dummies with the treatment are positive (with the only exception of psychology) even though heterogenous in magnitude and not always statistically different from zero. Thus, the average effect is not driven by the behavior of scholars belonging to a few peculiar fields.

VI.b Robustness checks

Our triple regression discontinuity model is an extended version of the regression discontinuity with multiple assignment variable proposed, among others, by Papay et al. (2011) and Papay et al. (2014). In particular, it is close to what the latter define as the 'Response-Surface RD'. These models depend heavily on a correct specification of the parametric functional form, as the gain in both efficiency and power resulting from multidimensionality comes at the expenses of lower flexibility.[30] Moreover, as for any full-parametric approach, the presence of (within-sector) outliers can bias the estimated coefficients, as all observations are assigned an equal weight irrespective of their distance from the cutoff.

To account for this latter issue, in our baseline specification, we exclude observations in the top decile and the bottom percentile of the distribution of distances from the cutoffs. In this section, we show that our main results are robust to adopting alternative sample restrictions. More specifically, we replicate our analysis varying the lower and the upper bounds of the distribution of distances from the cutoffs, thus progressively excluding candidates whose scores lie outside specified inter-percentile ranges (see Figure A4 in the Appendix). Results from this test show that considering a broader or narrower sample does not deliver significantly different estimates for the ITT effect. However, including observations in the far right tail of the distribution of the productivity indicators increases the noise in the sample and lowers the significance of the estimated coefficients. This finding is consistent with the fact that a candidate's publications have in principle no upper limit, whereas they cannot be less than zero. Thus, most of the outliers are located above the multidimensional cutoff.

To address possible concerns owing to the functional form assumed in our baseline estimation, we also replicate our analysis assuming a linear specification in the neighborhood of the thresholds. More precisely, we first normalize each running variable through dividing it by the corresponding cutoff value – thus accounting for between-field heterogeneity in candidates' average productivity

[30]Since we want to estimate the average treatment effect along the multidimensional borders, we cannot include a two- or three-dimensional spline since, by doing so, we would estimate a very local effect at the intersection of all cutoffs.

– and then select three different bandwidths, one for each running variable. Finally, we re-estimate Equation (2) on the sample of scholars whose productivity indicators lie within the resulting multi-dimensional joint bandwidth, assuming a linear specification.[31] Table A2 in the Appendix reports the result of this further robustness check and a comparison with our baseline results. The point estimates resulting from this local linear approach are very close in magnitude to those obtained assuming a second-degree polynomial form over the entire support. However, they are less precise, as standard errors are larger. Our preferred, fully parametric specification with field-specific interactions indeed allows us to better estimate the effect of complying with the two-out-of-three rule accounting for the between-field heterogeneity in the distribution of the productivity indicators. This goal is harder to achieve with a nonparametric approach within the neighborhood of the thresholds since we face a framework with multiple running variables and multiple field-specific cutoffs. The literature indeed lacks a procedure to compute optimal bandwidths in a similar context taking into account the wide across-field heterogeneity in the distribution of the running variables.

Additionally, we perform two placebo exercises to rule out the concern that our findings could be driven by systematic differences between candidates at the two sides of the cutoff rather than by a reaction to the treatment provision. First, we estimate our equations using the quantity of candidates' publications in each year before the 2012 NSQ as the dependent variable. In the case of any specification or sorting issues, our triple regression discontinuity model should also deliver non-zero results in the pre-treatment period. As shown in Table A4 in the Appendix, no discontinuity in terms of total publications and articles between treated and controls emerges when looking at each of the four years prior to the NSQ. Moreover, in Table II, we show that the marginal applicants in the two sides of the cutoff do not differ in terms of the aggregate quantity and quality of their publications or in the number of collaborations when these measures are computed over the whole 2009-2012 period. Second, we apply a perturbation to each field-specific threshold. We expect the magnitude of both our estimated first-stage and ITT coefficients to decline and the associate confidence intervals to broaden the farther we get from the original cutoff(s). Specifically,

[31] The bandwidths for the three productivity indicator are the MSE-optimal bandwidths computed separately for each running variable, following Calonico et al. (2014)

we reshuffle the cutoff values by adding a randomly generated error component $\epsilon \sim N(0, \sigma)$, which is defined as a percentage of the original field-specific cutoff.[32] The resulting perturbation, which we impose to lie within plus and minus the 100% of the original cutoff value, then has a different intensity depending on the standard deviation (σ) of the error. We then estimate the LATE from our baseline regression for increasing values of σ, replicating this exercise for 30 different draws from the ϵ distribution. We show in Figure A5 that the magnitude of the effect is the highest in the zero-perturbation case – that is, when using the true threshold values – and decreases in the variance of the perturbation. Taken together, the results from these two robustness tests confirm that our findings are not driven by any *ex ante* difference between candidates at the two sides of the multidimensional cutoff, thus lending important support to our identification strategy.

As a last robustness check, we test whether our results hold when using a different approach to deal with multiple applications. Since the rules of the NSQ allow candidates to apply for the qualification in different competition sectors, in our baseline specification, we consider for each candidate her 'best' application, that is, the one in which she scores the highest in terms of distance from the relevant thresholds. Here, we replicate our analysis considering for each applicant the indicators, the cutoffs, and the qualification outcome in the competition sector to which she already belongs as an assistant professor at the time of the application. Table A3 shows that the effect of achieving the qualification on the number of articles published between 2013 and 2016 is still positive and significant under this alternative specification. Coefficients are slightly lower in magnitude, consistent with the fact that, in this case, barely unsuccessful candidates might have succeeded in another competition sector. Therefore, a significant share of the candidates below the multidimensional cutoff are actually qualified and consequently exposed to promotion incentives, which makes the discontinuity in terms of post-call productivity smaller.

[32] We first generate the error $\epsilon \sim N(0, \sigma)$ and then draw from the ϵ distribution in order to assign a different perturbation to the cutoff value of each sector. We do this to account for the between-sector heterogeneity in each of the three productivity indicators.

VI.c Additional results

After analyzing the impact of passing the NSQ on the quantity of published items, we explore in this section whether it also affects other dimensions of the research activity of the academics in our sample. In particular, we investigate whether any significant discontinuity between (barely) qualified and non-qualified candidates emerges in terms of citations, publication quality and academic network size.

Citations. By replicating our baseline specification using the post-2012 citations received by each scholar as the dependent variable, we find that passing the qualification for the associate professorship also affects scholars' citations. The results in Table VI show that for papers published from 2013 to 2016, barely successful candidates receive on average 44 citations more than their barely unsuccessful colleagues (Column 1). This result can be attributed to both the increased number of publications of qualified scholars and the increase in the average number of citations *per* paper (Column 2). The probability of publishing an article with more than 50 citations (Column 3) or a non-cited article (Column 4) does not exhibit any jump, however.

Thus, scholars who are provided with higher promotion incentives in 2012 not only increase their publications but also manage to improve on another dimension that is taken into account in the qualification procedure: the number of citations received. This effect is in part simply driven by the increased research productivity of qualified scholars but could also reflect an augmented effort to promote and disseminate scientific works, greater visibility following a promotion, or an increase in the average publication quality. This latter aspect seems of particular importance and is therefore the next dimension on which we focus.

Average publication quality. The publication quality does not directly enter among the productivity indicators considered in the NSQ but could be indirectly affected by qualified scholars' incentives to maximize both citations and publications in a direction that is *a priori* ambiguous. On the one hand, publishing in better, more prestigious journals can increase a scholar's citations and H-index. On the other hand, there is a potential tension between the quantity and quality of publications, as submissions to prestigious journals are costly, especially in in terms of time, owing to the higher standards required and the more selective review processes. This trade-off could induce qualified scholars to sacrifice the quality dimension in order to minimize publication times and quickly increase their publication records.

We test these hypothesis by replicating our analysis using as dependent variables several alternative measures of a journal's quality and prestige. Specifically, we consider the *CiteScore*, the *Sjr* and *Snip* indexes, and the within-field *CiteScore* ranking – that is, a measure grouping journals according to their position in the field-specific distribution of the *CiteScore* index. According to the results reported in Table VII, Columns (2) to (5), no significant discontinuity in the average publication quality emerges, as all coefficients are not statistically different from zero. Additionally, we test whether the probability of publishing in a journal ranking in the top percentile of the CiteScore index (Column 1) or in a journal with no available measures of quality in the *Scopus* database (Column 6) changes discontinuously at the multidimensional threshold and find that this is not the case.

Hence, the documented increase in publications and citations by barely qualified scholars is not associated with a contemporaneous change in their average publication quality. Importantly, the large increase in the number of publications induced by the provision of promotion incentives does not appear to come at the expense of the average quality.

Co-author network. Finally, we study whether the outcome of the 2012 NSQ has any effect on the number of collaborations or on the size of scholars' co-authors network. In Table VIII, we report the estimated coefficients from our ITT and LATE equations, using as dependent variables i) the mean and the median number of authors per paper, ii) the probability of publishing a single-authored paper, and iii) the number of distinct co-authors. The three variables are computed for the 2013-2016 period. While the first two outcomes measure how each research paper is produced – that is, whether scholars tend to publish more or less coauthored works – the third proxies for the size of the academic network. We find suggestive evidence of a positive effect of achieving the associate professor qualification on scholars' co-authoring decisions, although the only significant (at the 10% level) coefficient is that for the median number of co-authors. Specifically, the estimated LATE in Column (2) shows that the median paper published by a barely qualified scholar has 2.2 more coauthors than that published by a barely unsuccessful scholar. These findings suggest that although some scholars might strategically expand their academic network in order to meet the thresholds for the full professor qualification more quickly, this behavior is not the main driver of the increase in productivity documented in the previous sections.

VII Distributional Effects and Competing Mechanisms

VII.a The distributional effects of promotion incentives

In this section, we dig deeper into the promotion incentives mechanism and explore whether scholars' reaction to achieving the NSQ is heterogeneous depending on the intensity of the incentives. Thus, we exploit across-individual differences in the distance between the productivity indicator and the full professor thresholds in 2012, the latter being the best estimate a candidate can have about the future thresholds she will face. This distance measures the size of the gap a scholar needs to fill in order to pass the (future) full professor threshold and therefore proxies for her chances of meeting the promotion thresholds in a relatively short time interval. Thus, we expect incentives to be low when the probability of obtaining the full professor qualification in the short or middle run is close to zero or to one, that is, when the gap that scholars have to fill is either too large or too small (see Lazear and Gibbs (2014)).

To reduce the dimensionality of the problem – and consistently with the dependent variable in our main regression (the number of publications post-2012) – we focus on the distance between the first bibliometric indicator – the number of articles published over the 2002-2012 period – and its (field-specific) full-professor cutoffs. Moreover, in order to account for the heterogeneity in pre-2012 research productivity and field characteristics, we normalize this distance dividing it by each candidate's number of publications as of 2012. The resulting index therefore measures the relative increase in publications that a candidate has to produce in order to reach the full professor cutoff. Thus, it constitutes a measure of the attainability of a further promotion in one of the subsequent calls of the NSQ.[33]

Figure VI.a reports point estimates and the associated 95% confidence intervals from regressing the number of post-2012 publications on our treatment, interacted with a categorical variable grouping observations according to the above-defined index. The estimated coefficients of the interaction terms show that the relationship between the increase in productivity and the distance from the promotion threshold is inverted-U shaped. The intention-to-treat effect is not significantly different from zero for those assistant professors who, in 2012, were either too close or too far from meeting the requirements for a full professor qualification. On the contrary, it is strongest for those in the middle of the relative-distance distribution. In order to fill the gap with the full professor cutoff, scholars in the third quartile would have to increase their stock of publications by approximately 25%.[34] This goal is realistic in a short- or middle-run horizon. Conversely, scholars in the last quintile would need to almost triple their stocks of publications, a target that is much more difficult to meet in a relatively short time interval. Importantly, this heterogeneity in the effect is not driven by across-quintiles differences in the probability of achieving the associate professor qualification in 2012 when complying with the two-out-of-three rule or in the pre-2012 research productivity of candidates. Both the first-stage coefficient and the ITT coefficient in the pre-2012 regression are indeed stable across the different quintiles (see Figure A6 in the Appendix). Furthermore, and consistent with the described mechanism, Figure VI.b documents that the probability of actually achieving the qualification for a full professorship by 2016 is heterogeneous across the above-defined quintiles. Candidates who were already very close to the full professor cutoff and those who increased their publication records the most after achieving the qualification are also those who are more likely to effectively achieve the qualification for a full professorship by the end of 2016. Conversely, candidates in the last two quintiles have a much lower likelihood to succeed

[33]More precisely, the index is defined as $dist_{i,1,s} = \frac{m^{full}_{1,s} - x_{i,1}}{x_{i,1}}$, where $m^{full}_{1,s}$ is the field-specific cutoff to overcome – in terms of the number of articles – to achieve the full professor qualification, while $x_{i,1}$ the professor's score in the same indicator.

[34]The stock of publications is computed in the ten years prior to the NSQ.

in the full professor NSQ in one of the following rounds.

Taken together, these two pieces of evidence show that candidates provided with the strongest incentives are also those who increase their post-2012 productivity the most, thus effectively improving their chances to succeed in one of the subsequent full professor qualification procedures. The effort induced by promotion incentives translates into an effectively higher probability of success. These results not only shed light on the distributional consequences of the promotion incentives induced by the qualification process but also lend important support for the promotion incentives mechanism. Most of the alternative explanation for our results – for instance, qualified scholars obtaining different teaching duties or easier access to research funds – would clash with the observed heterogeneity of the effect, depending on the variation in the intensity of the promotion incentives.

VII.b Promotion incentives *vs.* competing mechanisms

Results from our analysis document that scholars who attain the NSQ in 2012 increase the quantity of publications in the four years following the call. We interpret this finding as a response to the provision of promotion incentives: gaining the eligibility for an associate professorship 'unlocks' the possibility to achieve also the qualification for a full professorship. This, it incentivizes scholars to enrich their publication records so as to meet the requirements for the full professor qualification in the subsequent round.

Actually, there might be competing explanations for our result. For instance, obtaining the qualification could have a motivational effect, thus enhancing productivity, if it is perceived as a reward for past effort. Moreover, passing the qualification could induce substantial changes in scholars' daily life, as career advancements in academia are possibly associated with different teaching or bureaucratic duties, better access to research funds or broader networks. Still, this latter hypothesis seems inconsistent with the observed timing of the effect. Scholars' productivity begins rising immediately after the attainment of the mere eligibility for an associate professorship, rather than at the time of the actual promotion, which, for more than 75% of the qualified candidates, did not take place earlier than 2015. As a further exercise to address these concerns, we estimate our baseline equation for the sample of associate professors who apply for the full professor qualification in 2012. The NSQ indeed regulates both the access to associate and full professor positions, but candidates for this latter rank will have vanishing career incentives once the goal is achieved since no further advancements are possible. As a result, any effect detected in this sample of participant can hardly be reconciled with the promotion incentive mechanism proposed in this study.

The estimates reported in Table IX show that applicants who barely earn or barely miss the eligibility for the top academic position do not exhibit any significant difference in terms of later research productivity. This zero (or even negative) effect clashes with several alternative explanations for our main result. It shows that the effect of achieving the qualification is specific to the group of academics (tenured assistant professors) with further career prospects.

We also test whether the observed increase in publication might be due to the competition at the decentralized stage, where associate professorships are actually awarded. Achieving the NSQ might indeed incentivize qualified scholars to publish more in order to maximize their chances of obtaining an associate professor position as soon as a job vacancy opens rather than to meet the future full professor thresholds. Although data on scholars' promotion patterns suggest that there is limited across- and within-department competition, it could still be the case that the productivity jump is driven by fields with few vacancies and many qualified scholars competing for a position.[35]

Hence, we exploit the across-sector heterogeneity in the degree of internal competition for being hired as or promoted to an associate professor and test whether the effect of promotion incentives is actually stronger in sectors that feature more competition at the decentralized stage. Since we do not observe the actual number of vacancies but rather the equilibrium outcome, we use the ratio between the number assistant professors with an associate professor qualification and the number of existing associate professors in each field at the end of 2012 as a proxy for the degree of competition at the academic field level. This ratio indeed measures the ease of access to an associate

[35] Approximately two-thirds of eligible candidates in our sample obtained an associate professorship within three years from achieving the NSQ. Additionally, 97% of them obtained a promotion within the same university where they were employed in 2012.

professor position in a given competition sector, conditional on having achieved the qualification. Sectors in which there is a large mass of qualified candidates and few associate professors on staff are indeed likely to have lower turnover rates and therefore fewer vacancies, which make them relatively more competitive than those with a relatively low share of qualified scholars. Figure VII shows that the effect of passing the NSQ on the subsequent research productivity is very homogeneous across competition quintiles. This evidence, together with the distributional effects discussed in the previous section and with the zero-effect found on the sample of candidates to the full professor NSQ, strongly supports the fact that promotion incentives are the main mechanism at work.

VIII Conclusion

This paper studies the effectiveness of promotion incentives for high-skilled public employees. For a sample of 5,000 tenured assistant professors participating to the first round of the Italian NSQ – the centralized evaluation procedure awarding the eligibility for career advancements – we find that scholars exposed to a quasi-random provision of promotion incentives in 2012, owing to a success in the NSQ, increase the number of publications by almost 40% over the 2013-2016 period. Additionally, we find that the effect the incentives is the strongest for those scholars who are neither too far nor too close from the relevant future promotion thresholds. That is, promotion incentives are most effective when the promotion is "neither too hard to achieve, nor too easy" (Lazear and Gibbs, 2014, p.269). When exploring additional aspects of scholars' research activity, we find that qualified candidates receive more citations after the achievement of the NSQ and tend to expand the number of collaborations. The average publication quality – proxied by several measures of the journal's prestige – remains constant.

Several robustness tests and placebo exercises confirm the validity of our (triple) fuzzy regression discontinuity design. Further, consistent with the identification hypothesis according to which achieving the qualification affects productivity only through an increase in promotion incentives, we do not find a similar effect in the sample of associate professors applying for a full professorship.

Once the top ladder of the academic hierarchy is reached, achieving the qualification does not provide any further promotion incentives. Finally, we provide evidence that our results are not driven by possible changes in scholars' routine associated with a promotion or by the competition at the decentralized stage among qualified candidates.

These results shed light on a relatively unexplored topic in the existing personnel economics literature: the efficacy of promotion incentives in the public sector. This issue is particularly important as promotion incentives typically represent the main tool that public management can use to incentivize civil servants, as salaries in the public sector are typically rigid and make performance-pay schemes difficult to implement. Our findings show that promotion incentives, in the form of explicit and well-defined promotion thresholds, can effectively enhance public workers' productivity, especially when meeting the established targets requires a substantial but not excessive provision of effort. The policy implications of the analysis are relevant for countries in which the state personnel constitutes a significant share of the overall labor force and therefore, enhancing their performances can foster the overall productivity.

Lastly, our focus on a large, representative sample of academics adds further importance to the results of the analysis. According to the evidence we provide in this study, scholars strongly react to publication-based hiring and promotion schemes. Thus, the design of such mechanisms can represent an important instrument for policy-makers to promote the production of knowledge, the latter being a key factor for socio-economic development.

References

Ashraf, Nava, Oriana Bandiera, and Scott Lee, "Losing Prosociality in the Quest for Talent? Sorting, Selection, and Productivity in the Delivery of Public Services," Working Paper March 2018.

Bagues, Manuel, Mauro Sylos-Labini, and Natalia Zinovyeva, "Does the Gender Composition of Scientific Committees Matter?," *American Economic Review*, 2017, *107* (4), 1207–1238.

Baicker, Katherine and Mireille Jacobson, " Finders Keepers: Forfeiture Laws, Policing Incentives, and Local Budgets," *Journal of Public Economics*, 2007, *91* (11), 2113–2136.

Borjas, George J. and Kirk Doran, "Prizes and Productivity: How Winning the Fields Medal Affects Scientific Output," *Journal of Human Resources*, 2015, *50* (3), 728–758.

_ and _ , "Which Peers Matter? The Relative Impacts of Collaborators, Colleagues, and Competitors," *The Review of Economics and Statistics*, 2015, *97* (5), 1104–1117.

Calonico, Sebastian, Matias D. Cattaneo, and Rocio Titiunik, "Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs," *Econometrica*, 2014, *82* (6), 2295–2326.

_ , _ , Max H. Farrell, and Rocìo Titiunik, "Rdrobust: Software for Regression-Discontinuity Designs," *Stata Journal*, June 2017, *17* (2), 372–404.

Campbell, Dennis, "Nonfinancial Performance Measures and Promotion-Based Incentives," *Journal of Accounting Research*, 2008, *46* (2), 297–332.

Cattaneo, Matias D., Michael Jansson, and Xinwei Ma, "Simple Local Polynomial Density Estimators," Working Paper, University of Michigan 2017.

_ , _ , and _ , "Manipulation testing based on density discontinuity," *Stata Journal*, March 2018, *18* (1), 234–261.

Checchi, Daniele, Gianni De Fraja, and Stefano Verzillo, "Publish or Perish: An Analysis of the Academic Job Market in Italy," Discussion Papers 14/04, University of Nottingham, School of Economics July 2014.

Dal Bó, Ernesto, Frederico Finan, and Martín Rossi, "Strengthening State Capabilities: The Role of Financial Incentives in the Call to Public Service," *The Quarterly Journal of Economics*, 2013, *128* (3), 1169–1218.

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

_ , Michela Ponzo, and Vincenzo Scoppa, "Gender Differences in the Propensity to Apply for Promotion: Evidence from the Italian Scientific Qualification," *Oxford Economic Papers*, 2017, *69* (4), 986–1009.

Duflo, Esther, Rema Hanna, and Stephen Ryan, "Incentives Work: Getting Teachers to Come to School," *American Economic Review*, 2012, *102* (4), 1241–78.

Faria, J Ricardo and Peter McAdam, "Does Tenure Make Researchers Less Productive? The Case of the "Specialist"," School of Economics Discussion Papers 0514, School of Economics, University of Surrey 2014.

Finan, Frederico, Benjamin A. Olken, and Rohini Pande, "The Personnel Economics of the Developing State," *Handbook of Economic Field Experiments*, 2017, *2*, 467–514.

Friebel, Guido, Matthias Heinz, Miriam Krueger, and Nikolay Zubanov, "Team Incentives and Performance: Evidence from a Retail Chain," *American Economic Review*, 2017, *107* (8), 2168–2203.

Gaynor, Martin, James B. Rebitzer, and Lowell J. Taylor, "Physician Incentives in Health Maintenance Organizations," *Journal of Political Economy*, 2004, *112* (4), 915–931.

Gertler, Paul and Christel Vermeersch, "Using Performance Incentives to Improve Medical Care Productivity and Health Outcomes," NBER Working Papers 19046, National Bureau of Economic Research, Inc 2013.

Glewwe, Paul, Nauman Ilias, and Michael Kremer, "Teacher Incentives," *American Economic Journal: Applied Economics*, July 2010, *2* (3), 205–27.

Haeck, Catherine and Frank Verboven, "The Internal Economics of a University: Evidence from Personnel Data," *Journal of Labor Economics*, 2012, *30* (3), 591 – 626.

Iaria, Alessandro, Carlo Schwarz, and Fabian Waldinger, "Frontier Knowledge and Scientific Production: Evidence from the Collapse of International Science," *The Quarterly Journal of Economics*, 2018, *133* (2), 927–991.

Karachiwalla, Naureen and Albert Park, "Promotion Incentives in the Public Sector: Evidence from Chinese Schools," *Journal of Public Economics*, 2017, *146* (C), 109–128.

Khan, Adnan Q., Asim I. Khwaja, and Benjamin A. Olken, "Tax Farming Redux: Experimental Evidence on Performance Pay for Tax Collectors," *The Quarterly Journal of Economics*, 2016, *131* (1), 219–271.

_ **, Asim Ijaz Khwaja, and Benjamin A. Olken**, "Making Moves Matter: Experimental Evidence on Incentivizing Bureaucrats through Performance-Based Postings," *American Economic Review*, 2018, *forthcoming.*

Kwon, Illoong, "Incentives, Wages, and Promotions: Theory and Evidence," *The RAND Journal of Economics*, 2006, *37* (1), 100–120.

Lavy, Victor, "Evaluating the Effect of Teachers' Group Performance Incentives on Pupil Achievement," *Journal of Political Economy*, 2002, *110* (6), 1286–1317.

Lazear, Edward P., "Performance Pay and Productivity," *American Economic Review*, December 2000, *90* (5), 1346–1361.

_ **and Mike Gibbs**, *Personnel Economics in Practice*, John Wiley & Sons, 2014.

_ **and Paul Oyer**, "Personnel Economics," in Robert Gibbons and John Roberts, eds., *The Handbook of Organizational Economics*, Princeton University Press, 2013.

McCrary, Justin, "Manipulation of the Running Variable in the Regression Discontinuity Design: A Density Test," *Journal of Econometrics*, 2008, *142* (2), 698–714.

Muralidharan, Karthik and Venkatesh Sundararaman, "Teacher Performance Pay: Experimental Evidence from India," *Journal of Political Economy*, 2011, *119* (1), 39 – 77.

OECD, *Government at a Glance 2017*, OECD Publishing, 2017.

Olken, Benjamin A., Junko Onishi, and Susan Wong, "Should Aid Reward Performance? Evidence from a Field Experiment on Health and Education in Indonesia," *American Economic Journal: Applied Economics*, 2014, *6* (4), 1–34.

Papay, John P., Richard J. Murnane, and John B. Willett, "Extending the Regression-Discontinuity Approach to Multiple Assignment Variables," *Journal of Econometrics*, 2011, *161* (2), 203–207.

_ **,** _ **, and** _ , "High-School Exit Examinations and the Schooling Decisions of Teenagers: Evidence From Regression-Discontinuity Approaches," *Journal of Research on Educational Effectiveness*, 2014, *7* (1), 1–27.

Shearer, Bruce, "Piece Rates, Fixed Wages and Incentives: Evidence from a Field Experiment," *The Review of Economic Studies*, 2004, *71* (2), 513–534.

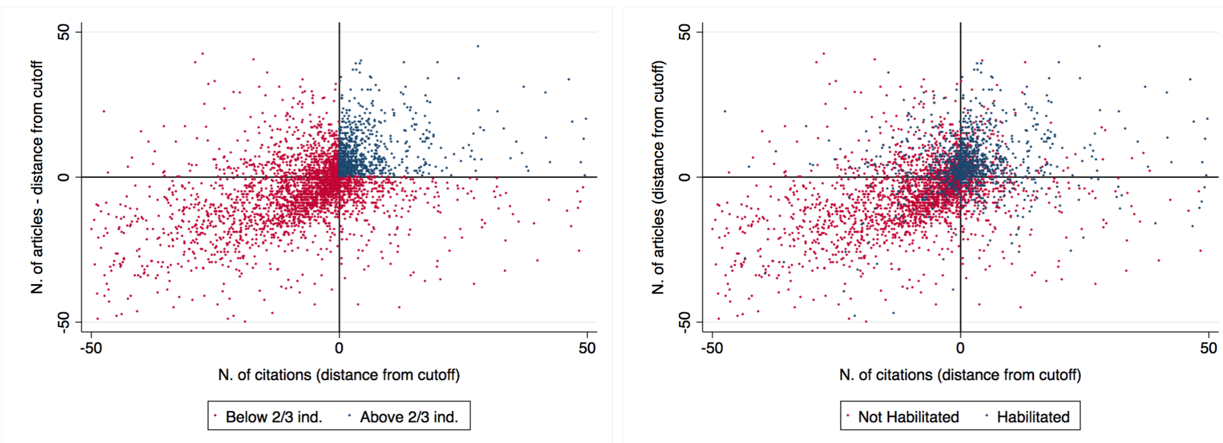
Waldinger, Fabian, "Peer Effects in Science: Evidence from the Dismissal of Scientists in Nazi Germany," *Review of Economic Studies*, 2012, *79* (2), 838–861.

_ , "Bombs, Brains, and Science: The Role of Human and Physical Capital for the Creation of Scientific Knowledge," *The Review of Economics and Statistics*, December 2016, *98* (5), 811–831.

Zinovyeva, Natalia and Manuel Bagues, "The Role of Connections in Academic Promotions," *American Economic Journal: Applied Economics*, 2015, *7* (2), 264–292.

Figures

Figure I: Distribution of candidates with respect to distance from cutoffs

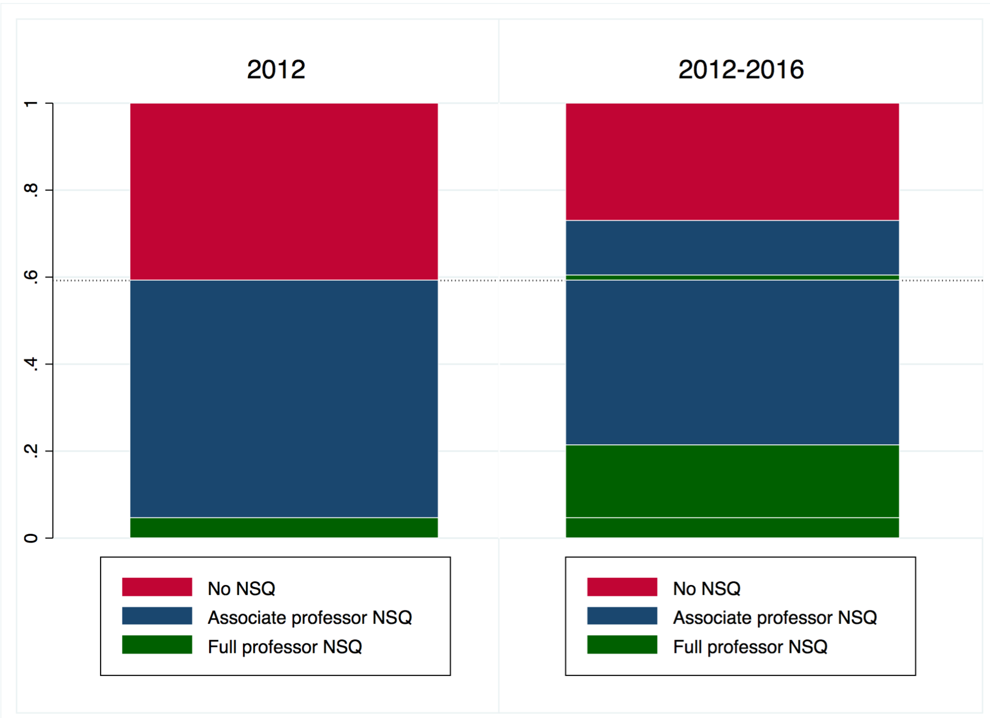


a. Compliance with the two-out-of-three rule

b. Qualified

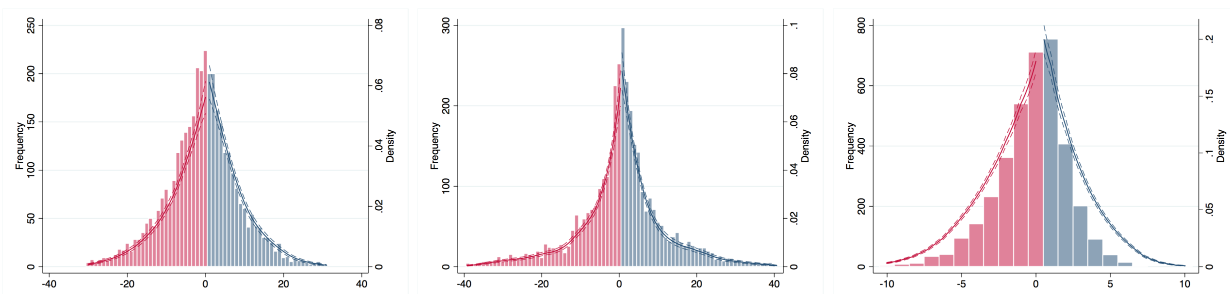
Notes. This figure depicts the distribution of candidates for the associate professor NSQ in 2012 in bibliometric fields, depending on the distances between the first two bibliometric indicators and the corresponding cutoffs. The distance between the number of articles and the cutoff is on the y-axis, while the distance between the number of citations and the cutoff is on the x-axis. In both panels, we consider only applicants whose H-Index is below the cutoff. Therefore, circles in the upper-right quadrant correspond to candidates complying with the two-out-of-three rule. Blue circles indicate qualified candidates, while red circles indicate unsuccessful candidates. Within each academic field, we exclude observations in the top 10% and the bottom 1% of the distribution of the running variables. We also exclude fields with more than 90% of successful candidates and those with less than 30 observations.

Figure II: NSQ trajectories 2012-2016



Notes. This figure depicts the share of successful and unsuccessful assistant professors applying to the NSQ between 2012 and 2016 in bibliometric fields. The left bar reports the share of candidates who were not qualified (red), qualified for an associate professorship (blue) and qualified directly for a full professorship (green). For each of the three groups, in the right bar we report the corresponding shares as of the end of 2016, that is, after the 2013 and 2016 calls of the NSQ.

Figure III: Frequency distribution and manipulation test



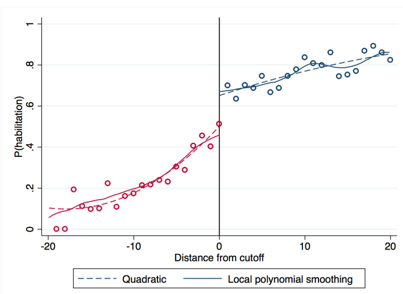
a. N. of articles

b. N. of citations

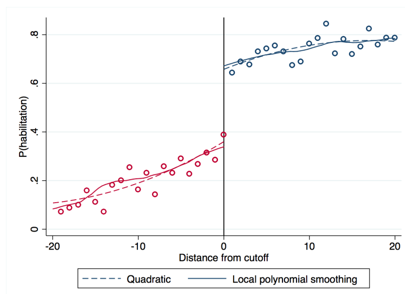
c. H-index

Notes. This figure depicts the frequency distribution and the estimated kernel density distribution of candidates for the associate professor NSQ in 2012 in bibliometric fields depending on their distance from each of the three bibliometric cutoffs. The frequency distributions in Panels A, B, and C are constructed within the intervals $[-30, 30]$, $[-40, 40]$, $[-10, 10]$ for indicators 1, 2 and 3, respectively. In all panels, the bin width is equal to 1. The kernel density is estimated following McCrary (2008).

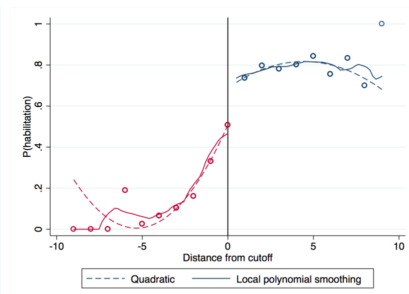
Figure IV: Discontinuity effect on the probability of success in the NSQ



a. N. of articles



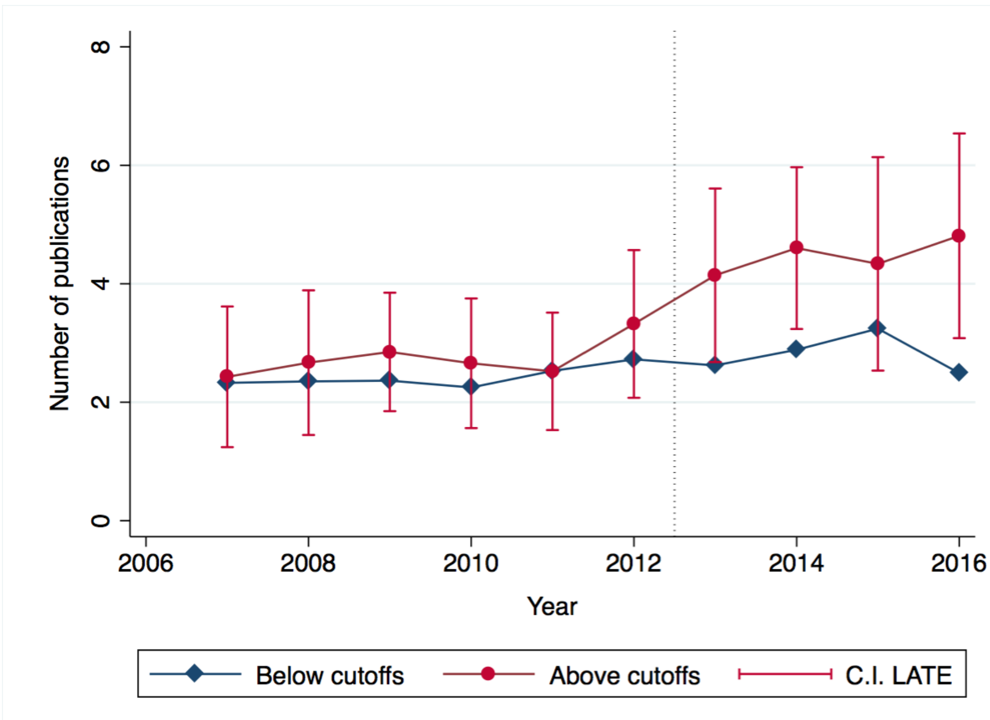
b. N. of citations



c. H-index

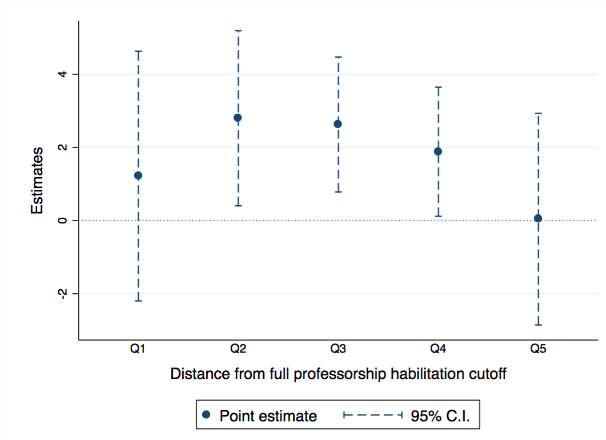
Notes. This figure depicts the discontinuous jumps in the probability of achieving the qualification arising when each of the three indicators crosses the relevant cutoff value. The sample includes the candidates for the associate professor NSQ in 2012 in bibliometric fields. Each circle represents the average probability of achieving the NSQ within each 1-unit bin. The running variables for the three indicators are defined as the distance from the field-specific median, thus centered at zero. The dependent variable in the quadratic and local polynomial smoothing regression is an indicator that equals one when a candidate obtains the qualification. Both the quadratic and the local polynomial smoothing regressions are estimated within a $[-20, 20]$ interval of the distance from the thresholds for the first two indicators (Panel A and B), and within a $[-10, 10]$ interval of the distance from the threshold of the third indicator (Panel C). Within each academic field, we exclude observations in the top 10% and the bottom 1% of the distribution of the distances. We also exclude the fields with more than 90% successful candidates and those with fewer than 30 observations.

Figure V: RD estimates of achieving the NSQ on the number of publications

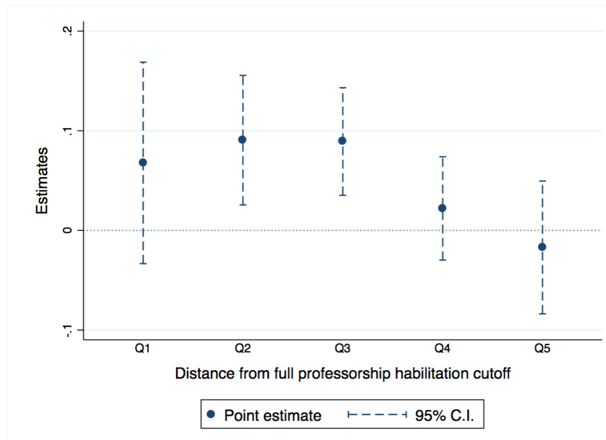


NOTES. This figure depicts the OLS coefficients of the indicator for the compliance with the two-out-of-three rule on the quantity of publications estimated separately for each year over the 2007-2016 period. The sample includes the candidates for the associate professor NSQ in 2012 in bibliometric fields. Each blue dot corresponds to a weighted average of the field-specific dummies included in the regression for the specified year; each red dot corresponds to the LATE coefficient plus the 'weighted' constant term. Thus, the distance between the two dots represents the LATE for each regression. The LATE is the 2SLS coefficient of the indicator that equals one when a candidate achieves the qualification. The independent variable in all regressions is instrumented by the indicator that equals one when a candidate complies with the two-out-of-three rule. First-stage estimates are reported in Column (7) of Table III. The dependent variables is the total number of papers (including articles, conference papers, reviews and other items) in the specified year. In each regression, within each academic field, we exclude observations in the top 10% and the bottom 1% of the distribution of the distances. We also exclude the fields with more than 90% successful candidates and those with fewer than 30 observations. All regressions are estimated using a polynomial (quadratic) specification over the entire support.

Figure VI: Effect heterogeneity: distance from the full professor thresholds



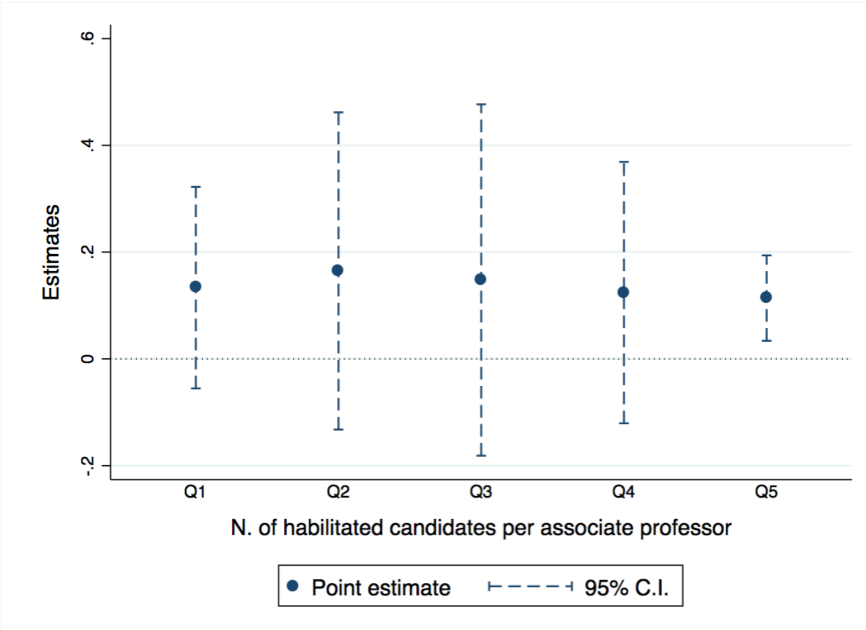
a. Number of publications (2013-2016)



b. Probability of achieving the full professor NSQ (2013-2016)

NOTES. This figure depicts the OLS coefficients of the interaction between the indicator for the compliance with the two-out-of-three rule and a set of indicators summarizing the heterogeneity in the distance from the full-professor thresholds on the quantity of publications over the 2013-2016 period. The sample includes the candidates for the associate professor NSQ in 2012 in bibliometric fields. The blue circles correspond to the coefficient of an interaction term formed by multiplying an indicator that equals one when a candidate complies with the two-out-of-three rule with a set of indicators that equal 1 if a candidate belongs to the specified quintile of the distribution of the distance from the full professor threshold. The distance from the full professor threshold is expressed as a percentage of the initial stock of articles, thus defined as $dist_{i,1,s} = \frac{m^{full}_{1,s} - x_{i,1}}{x_{i,1}}$, where $m^{full}_{1,s}$ is the field-specific cutoff for the full professor NSQ and $x_{i,1}$ is the professor's score in the same indicator. The dependent variable is the total number of papers (including articles, conference papers, reviews and other items) published during the 2013-2016 period. In each regression, within each academic field, we exclude observations in the top 10% and the bottom 1% of the distribution of the distances. We also exclude the fields with more than 90% successful candidates and those with fewer than 30 observations. All regressions are estimated using a polynomial (quadratic) specification over the entire support.

Figure VII: Effect heterogeneity: degree of internal competition



Notes. This figure depicts the OLS coefficients of the interaction between the indicator for compliance with the two-out-of-three rule and a set of indicators summarizing the heterogeneity in the degree of competition in the decentralized stage on the quantity of publications over the 2013-2016 period. The sample includes the candidates for the associate professor NSQ in 2012 in bibliometric fields. The blue circles correspond to the coefficient of an interaction term formed by multiplying an indicator that equals one when a candidate complies with the two-out-of-three rule with a set of indicators that equal one if a candidate belongs to the specified quintile of the distribution of the degree of competition in the decentralized stage. The degree of competition in the decentralized stage is defined as the ratio between the number of assistant professors with an associate professor NSQ and the number of associate professors already employed in each field at the end of 2012. The dependent variable is the total number of papers (including articles, conference papers, reviews and other items) published during the 2013-2016 period. In each regression, within each academic field we exclude observations in the top 10% and the bottom 1% of the distribution of the distances. We also exclude the fields with more than 90% successful candidates and the ones with less than 30 observations. All regressions are estimated using a polynomial (quadratic) specification over the entire support.

Tables

Table I: Descriptive Statistics

Panel A: applicants' characteristics				
	Mean		Sd	
Academic age	15.87		7.65	
Female	0.42		0.49	
N. of applications	1.26		0.65	
Qualified	0.58		0.49	
Above 2/3 medians	0.66		0.47	
Above median (n. of articles)	0.62		0.49	
Above median (n. of citations)	0.68		0.47	
Above median (h-index)	0.58		0.49	
Distance from field median (n. of articles)	4.43		18.26	
Distance from field median (n. of citations)	9.62		38.36	
Distance from field median (h-index)	0.92		3.55	
N. of scholars	4920			
Panel B: publications 2008-2016				
	2008-2011		2013-2016	
	Mean	Sd	Mean	Sd
Number of publications	12.40	10.37	17.01	20.41
Number of articles	8.67	7.42	12.75	17.82
Number of conference papers	2.29	5.37	2.31	5.72
Number of reviews	0.68	1.55	0.92	2.09
N. of scholars	4920			
Panel C: characteristics of the publications 2008-2016				
	2008-2011		2013-2016	
	Mean	Sd	Mean	Sd
Top 1% journal	0.03	0.16	0.02	0.15
CiteScore (percentile)	72.18	24.79	75.44	23.19
CiteScore	2.69	2.26	2.80	2.13
Sjr	1.48	1.68	1.52	1.58
Snip	1.27	0.94	1.31	0.91
Journal unlisted	0.14	0.35	0.13	0.34
N. of citations received	17.19	38.81	5.34	15.00
Single-authored	0.02	0.15	0.02	0.12
N. of authors	8.27	12.03	10.82	17.09
N. of publications	43810		83670	

NOTES. This table reports the baseline characteristics of candidates for the associate professor NSQ in 2012 in bibliometric fields together with the summary statistics about their research activity over the period 2008-2016. The unit of analysis is the single candidate in Panels (A) and (B), and the single publication in Panel (C). The variables 'Top 1% journal', 'Journal unlisted', and 'Single-authored' are binary indicators which take the value of one if the publication appears in journals scoring in the top 1% of the distribution of the 2012 CiteScore journal percentile index, it appears in journals not classified in the *Scopus* database, or it has a single author, respectively. In all Panels, within each academic field we exclude observations in the top 10% and the bottom 1% of the distribution of the distances. We also exclude the fields with more than 90% successful candidates and those with fewer than 30 observations.

Table II: Continuity test (2009-2011 measures)

Candidates for associate professor	(1) N. of publications	(2) Zero publications	(3) % Unlisted	(4) Top 5%	(5) CiteScore (pct)	(6) N. Coauthors (mean)	(7) Single-author	(8) Network size
Above 2/3 cutoffs	0.392	-0.002	0.028	0.173	0.010	0.310	-0.013	1.372
	(0.467)	(0.007)	(0.034)	(1.143)	(0.010)	(0.315)	(0.025)	(3.608)
Mean Dep. Var.	11.658	0.026	0.580	70.999	0.122	6.860	0.143	41.859
Standard dev.	9.598	0.159	0.494	15.860	0.196	5.766	0.351	53.641
N. of clusters	82	82	82	82	82	82	82	82
Observations	4920	4920	4755	4755	4792	4763	4861	4785

Notes. This table reports the OLS coefficients of overcoming two out of the three field-specific bibliometric cutoffs on the quantity of publications, their quality and the number of collaborations computed over the period 2009-2011. The sample includes the candidates for the associate professor NSQ in 2012 in bibliometric fields. The dependent variables are: the total number of papers – including articles, conference papers, reviews and other items– (Column 1); an indicator that equals one when a candidate does not publish any paper (Column 2); the share of publications in journals not classified in the *Scopus* database (Column 3); the share of articles published in journals scoring in the top 5% according to the 2012 *CiteScore journal percentile* (Column 4); the average *CiteScore journal percentile* (Column 5); the average number of co-authors per publication (Column 6); the share of single-authored publications (Column 7); and the total number of distinct co-authors (Column 8). In Column (6) the sample includes only scholars with at least one record in the *Scopus* database during the period 2013-2016. In Columns (3) to (5) the sample is further limited to scholars with at least one publication in a journal classified in the *Scopus* database (with a non-missing score) during the period 2009-2011. Within each academic field, we exclude observations in the top 10% and the bottom 1% of the distribution of the distances. We also exclude the fields with more than 90% successful candidates and those with fewer than 30 observations. All regressions are estimated using a polynomial (quadratic) specification over the entire support.
Standard errors, clustered at the academic field level, in parentheses. *** p< 0.01, ** p<0.05, and *p<0.10.

Table III: First stage

Candidates for Associate professor							
	Single RD (Articles)		Single RD (Citations)		Single RD (H-Index)		Triple RD
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Quadratic	LLR	Quadratic	LLR	Quadratic	LLR	Quadratic
Above cutoff	0.131***	0.122***	0.210***	0.161***	0.153***	0.162**	
	(0.040)	(0.046)	(0.041)	(0.057)	(0.046)	(0.066)	
Above 2/3 cutoffs							0.306***
							(0.043)
Academic field FE	Yes	No	Yes	No	Yes	No	Yes
Field specific interactions	Yes	No	Yes	No	Yes	No	Yes
Mean dep. var.	0.577	0.573	0.577	0.536	0.577	0.564	0.577
Sd dep. var.	0.494	0.495	0.494	0.499	0.494	0.496	0.494
BW (MSE)		0.396		0.364		0.304	
N of clusters	82	89	82	89	82	89	82
Observations	4920	2752	4920	1798	4920	3034	4920

NOTES. This table reports the OLS coefficients of overcoming the field-specific bibliometric cutoffs on the outcome of the 2012 NSQ. The sample includes the candidates for the associate professor NSQ in 2012 in bibliometric fields. In all columns, the dependent variable is an indicator that equals one when a candidate obtains the qualification, and zero otherwise. In Columns (1) to (6), the main independent variable is an indicator that equals one when a candidate overcomes the relevant cutoff for either the number of articles (Columns 1 and 2), the number of citations (Columns 3 and 4) and the h-index (Columns 5 and 6). In Column (7), the main independent variable is an indicator that equals one when a candidate complies with the two-out-of-three rule, that is, when her scores in at least two indicators are above the relevant cutoffs. Regressions in Columns (1), (3), (5) and (6) are estimated using a field-specific polynomial (quadratic) specification over the entire support after excluding observations in the top 10% and the bottom 1% of the within-field distribution of the distances from each cutoff. In Columns (2), (4) and (6), we replicate the estimates in (1), (3), and (5), performing local linear regressions (LLR) within the MSE-optimal bandwidths computed following Calonico et al. (2014) – using the companion Stata package described in Calonico et al. (2017) – after normalizing each distance from the cutoff by dividing it by the cutoff itself. In all specifications, we exclude the fields with more than 90% successful candidates and those with fewer than 30 observations.
Standard errors, clustered at the academic field level, in parentheses. *** p< 0.01, ** p<0.05, and *p<0.10.

Table IV: The effect of achieving the NSQ on the number of papers published

Candidates for Associate professor				
	Publications	Articles	Conf. Papers	Reviews
	(1)	(2)	(3)	(4)
ITT	2.003***	1.254**	0.540	0.198*
	(0.701)	(0.477)	(0.336)	(0.102)
LATE	6.557***	4.105***	1.769*	0.648**
	(2.159)	(1.485)	(0.956)	(0.290)
Mean Dep. Var.	17.006	12.747	2.313	0.916
Standard dev.	20.410	17.820	5.719	2.088
N. of clusters	82	82	82	82
Observations	4920	4920	4920	4920

Notes. This table reports the ITT and the LATE of achieving the qualification on the quantity of publications over the 2013-2016 period. The sample includes the candidates for the associate professor NSQ in 2012 in bibliometric fields. The dependent variable in Column (1) is the total number of papers (including articles, conference papers, reviews and other items) published during the 2013-2016 period; the dependent variables in Columns (2), (3) and (4) are the total number of articles, conference papers and reviews published during the 2013-2016 period, respectively. The ITT is the OLS coefficient of the indicator that equals one when a candidate complies with the two-out-of-three rule; the LATE is the 2SLS coefficient of the indicator that equals one when a candidate obtains the qualification. The latter is instrumented by the indicator that equals one when a candidate complies with the two-out-of-three rule. First-stage estimates are reported in Column (7) of Table III. Within each academic field, we exclude observations in the top 10% and the bottom 1% of the distribution of the distances. We also exclude the fields with more than 90% successful candidates and those with fewer than 30 observations. All regressions are estimated using a polynomial (quadratic) specification over the entire support.
Standard errors, clustered at the academic field level, in parentheses. *** $p< 0.01$, ** $p<0.05$, and *$p<0.10$.

Table V: Effect heterogeneity: gender

Dependent variable: Number of publications		
	First stage	LATE
	(1)	(2)
female	0.005	-1.148
	(0.020)	(0.875)
above 2/3 cutoffs	0.329***	
	(0.046)	
above 2/3 cutoffs × female	-0.057*	
	(0.030)	
qualified		6.742***
		(2.253)
qualified × female		-0.684
		(1.563)
Mean Dep. Var.	0.577	17.006
Standard dev.	0.494	20.410
N. of clusters	82	82
Observations	4920	4920

Notes. This table reports the OLS coefficient of the indicator for the compliance with the two-out-of-three rule and its interaction with gender on the outcome of the 2012 NSQ (Column 1) and on the quantity of publications over the period 2013-2016 (Column 2). The sample includes the candidates for the associate professor NSQ in 2012 in bibliometric fields. The dependent variable in Column (1) is an indicator that equals one when a candidate obtains the qualification, and zero otherwise; the dependent variable in Column (2) is the total number of papers (including articles, conference papers, reviews and other items) published during the 2013-2016 period. In both columns, the main independent variables are an indicator that equals one when a candidate complies with the two-out-of-three rule, an indicator that equals one in case of a female candidate and an interaction term formed by multiplying the two indicators. Within each academic field, we exclude observations in the top 10% and the bottom 1% of the distribution of the distances. We also exclude the fields with more than 90% successful candidates and those with fewer than 30 observations. All regressions are estimated using a polynomial (quadratic) specification over the entire support.
Standard errors, clustered at the academic field level, in parentheses. *** p< 0.01, ** p<0.05, and *p<0.10.

Table VI: The effect of achieving the NSQ on the number of citations received

Candidates for Associate professor				
	Total citations	Mean citations	% Cit. ≥ 50	% Zero cit.
ITT	13.526**	0.519*	0.002	0.012
	(6.372)	(0.273)	(0.002)	(0.012)
LATE	43.934**	1.687**	0.006	0.039
	(18.087)	(0.748)	(0.004)	(0.034)
Mean Dep. Var.	92.371	4.417	0.005	0.356
Standard dev.	200.847	4.184	0.028	0.211
N. of clusters	82	82	82	82
Observations	4838	4838	4838	4838

Notes. This table reports the ITT and the LATE of achieving the qualification on the citations received by papers published over the 2013-2016 period. The sample includes the candidates for the associate professor NSQ in 2012 in bibliometric fields with at least one record in the *Scopus* database over the 2013-2016 period. The dependent variables in Columns (1) and (2) are the total and the average number of citations received by papers published during the 2013-2016 period, respectively. The dependent variable in Columns (3) and (4) are the share of papers published during the 2013-2016 period with at least 50 citations and with zero citations, respectively. The ITT is the OLS coefficient of the indicator that equals one when a candidate complies with the two-out-of-three rule; the LATE is the 2SLS coefficient of the indicator that equals one when a candidate obtains the qualification. The latter is instrumented by the indicator that equals one when a candidate complies with the two-out-of-three rule. First-stage estimates are reported in Column (7) of Table III. Within each academic field, we exclude observations in the top 10% and the bottom 1% of the distribution of the distances. We also exclude the fields with more than 90% successful candidates and those with fewer than 30 observations. All regressions are estimated using a polynomial (quadratic) specification over the entire support.
Standard errors, clustered at the academic field level, in parentheses. *** p< 0.01, ** p<0.05, and *p<0.10.

Table VII: The effect of achieving the NSQ on the quality of publications

Candidates for Associate professor						
	Top 5%	CiteScore (pct)	CiteScore	Sjr	Snip	% Unlisted
	(1)	(2)	(3)	(4)	(5)	(6)
ITT	0.038	-0.110	0.063	0.052	0.040	0.000
	(0.030)	(1.087)	(0.069)	(0.059)	(0.032)	(0.011)
LATE	0.126	-0.360	0.206	0.170	0.130	0.001
	(0.083)	(3.065)	(0.192)	(0.168)	(0.086)	(0.029)
Mean Dep. Var.	0.655	72.403	2.509	1.336	1.239	0.126
Standard dev.	0.475	14.773	1.343	0.843	0.456	0.180
N. of clusters	82	82	82	82	82	82
Observations	4809	4809	4809	4808	4809	4838

Notes. This table reports the ITT and the LATE of achieving the qualification on the quality of publications over the 2013-2016 period. The sample includes the candidates for the associate professor NSQ in 2012 in bibliometric fields. The dependent variable in Column (1) is the share of articles published in journals scoring in the top 5% according to the 2015 *CiteScore journal percentile* during the 2013-2016 period; the dependent variables in Columns (2), (3), (4) and (5) are the average *CiteScore journal percentile*, the average *CiteScore* index, the average *Sjr* index, the average *Snip* index of papers published during the 2013-2016 period, respectively; the dependent variable In Column (6) is the share of publications in journals not classified in the *Scopus* database during the period 2013-2016. The ITT is the OLS coefficient of the indicator that equals one when a candidate complies with the two-out-of-three rule; the LATE is the 2SLS coefficient of the indicator that equals one when a candidate obtains the qualification. The latter is instrumented by the indicator that equals one when a candidate complies with the two-out-of-three rule. First-stage estimates are reported in Column (7) of Table III. In Column (6) the sample includes only scholars with at least one record in the *Scopus* database during the period 2013-2016. In Columns (2) to (5) the sample is further limited to scholars with at least one publication in a journal classified in the *Scopus* database of journals (with a non-missing score) during the period 2013-2016. Within each academic field, we exclude observations in the top 10% and the bottom 1% of the distribution of the distances. We also exclude the fields with more than 90% successful candidates and the ones with fewer than 30 observations. All regressions are estimated using a polynomial (quadratic) specification over the entire support.
Standard errors, clustered at the academic field level, in parentheses. *** $p< 0.01$, ** $p<0.05$, and *$p<0.10$.

Table VIII: The effect of achieving the NSQ on the number of collaborations

Candidates for Associate professor				
	(1)	(2)	(3)	(4)
	N. Coauthors (mean)	N. Coauthors (median)	Single-author	Network size
ITT	0.241	0.686	0.016	1.748
	(0.459)	(0.419)	(0.029)	(5.334)
LATE	0.785	2.236*	0.053	5.722
	(1.309)	(1.292)	(0.082)	(15.061)
Mean Dep. Var.	8.185	7.255	0.137	62.016
Standard dev.	8.245	8.217	0.344	84.882
N. of clusters	82	82	82	82
Observations	4801	4801	4801	4806

Notes. This table reports the ITT and the LATE of achieving the qualification on the number of co-authorships over the 2013-2016 period. The sample includes the candidates for the associate professor NSQ in 2012 in bibliometric fields with at least one record in the *Scopus* database over the 2013-2016 period. The dependent variables are the average number of coauthors per publication (Column 1), the maximum number of co-authors per publication (Column 2), the share of single-authored publications (Column 3) and the total number of distinct coauthors (Column 4). The ITT is the OLS coefficient of the indicator that equals one when a candidate complies with the two-out-of-three rule; the LATE is the 2SLS coefficient of the indicator that equals one when a candidate obtains the qualification. The latter is instrumented by the indicator that equals one when a candidate complies with the two-out-of-three rule. First-stage estimates are reported in Column (7) of Table III. Within each academic field, we exclude observations in the top 10% and the bottom 1% of the distribution of the distances. We also exclude the fields with more than 90% successful candidates and the ones with less than 30 observations. All regressions are estimated using a polynomial (quadratic) specification over the entire support.
Standard errors, clustered at the academic field level, in parentheses. *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.10$.

Table IX: The effect of achieving the full professor NSQ on the number of papers published

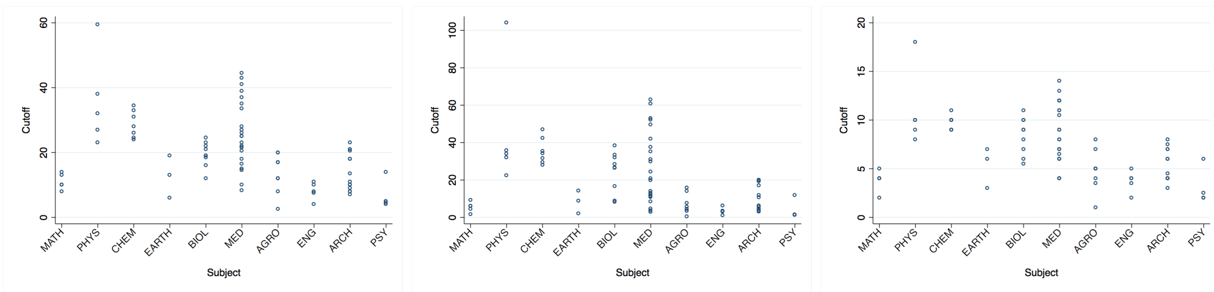
Candidates for Full professor				
	Publications	Articles	Conf. Papers	Reviews
	(1)	(2)	(3)	(4)
ITT	-0.349	-0.251	0.195	-0.105
	(1.195)	(1.009)	(0.221)	(0.189)
LATE	-0.836	-0.601	0.468	-0.251
	(2.408)	(2.037)	(0.467)	(0.382)
Mean Dep. Var.	22.592	16.776	2.727	1.431
Standard dev.	24.987	21.383	6.982	2.758
N. of clusters	51	51	51	51
Observations	2746	2746	2746	2746

Notes. This table reports the ITT and the LATE of achieving the qualification on the quantity of publications over the 2013-2016 period. The sample includes the candidates for the associate professor NSQ in 2012 in bibliometric fields. The dependent variable in Column (1) is the total number of papers (including articles, conference papers, reviews and other items) published during the 2013-2016 period; the dependent variables in Columns (2), (3) and (4) are the total number of articles, conference papers and reviews published during the 2013-2016 period, respectively. The ITT is the OLS coefficient of the indicator that equals one when a candidate complies with the two-out-of-three rule; the LATE is the 2SLS coefficient of the indicator that equals one when a candidate obtains the qualification. The latter is instrumented by the indicator that equals one when a candidate complies with the two-out-of-three rule. First-stage estimates are reported in Column (7) of Table III. Within each academic field, we exclude observations in the top 10% and the bottom 1% of the distribution of the distances. We also exclude the fields with more than 90% successful candidates and those with fewer than 30 observations. All regressions are estimated using a polynomial (quadratic) specification over the entire support.
Standard errors, clustered at the academic field level, in parentheses. *** $p < 0.01$, ** $p<0.05$, and *$p<0.10$.

Appendix

A.1 Figures

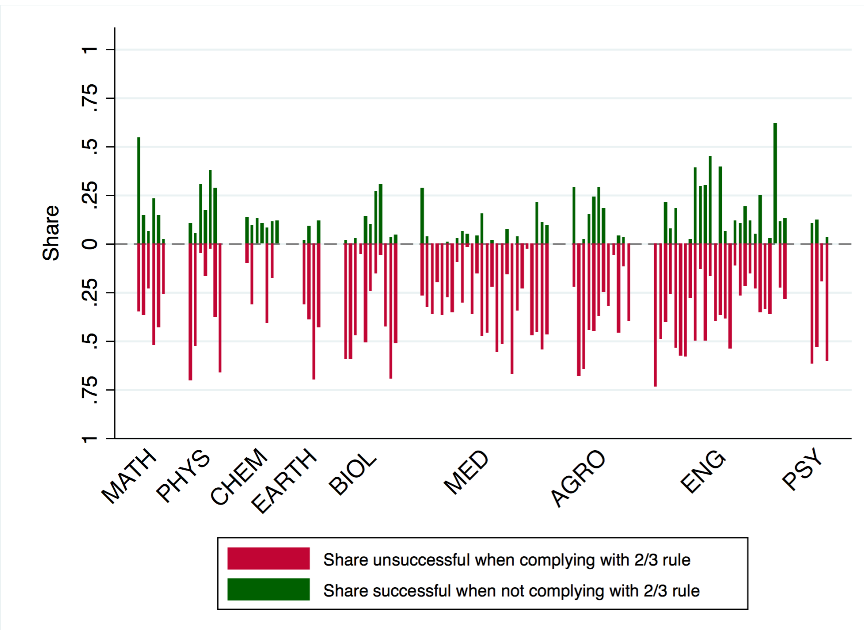
Figure A1: Heterogeneity in field-specific cutoffs



a. Number of articles **b.** Number of citations **c.** H-index

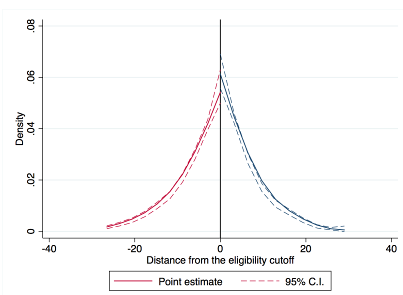
NOTES. This figure depicts the value of the cutoffs for each bibliometric indicator and academic field. Each circle corresponds to an academic field, grouped according to its main subject. We exclude the fields with more than 90% successful candidates and those with fewer than 30 observations.
Legend: MATH=Mathematics; PHYS=Physics; CHEM=Chemistry; EARTH=Earth Sciences; BIOL= Biology; MED=Health Sciences; AGRO=Agronomy and Veterinary; ENG=Engineering; ARCH=Architecture; PSY=Psychology.

Figure A2: Compliance with the two-out-of-three rule

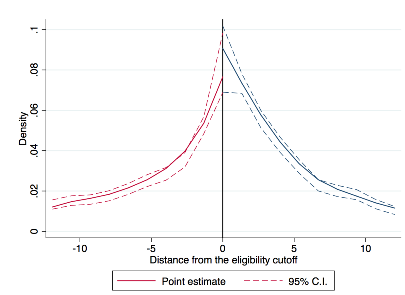


NOTES. This figure depicts the share of candidates for the associate professor NSQ in 2012 in bibliometric fields, depending on the compliance with the two-out-of-three rule and the outcome of the 2012 NSQ. Each bar corresponds to an academic field, grouped according to its main subject. The length of each red (green) bar indicates the share of candidates who did not obtain (obtained) the qualification in the 2012 NSQ, even when complying (not complying) with the two-out-of-three rule. We exclude the fields with more than 90% successful candidates and those with fewer than 30 observations.
Legend: MATH=Mathematics; PHYS=Physics; CHEM=Chemistry; EARTH=Earth Sciences; BIOL= Biology; MED=Health Sciences; AGRO=Agronomy and Veterinary; ENG=Engineering; ARCH=Architecture; PSY=Psychology.

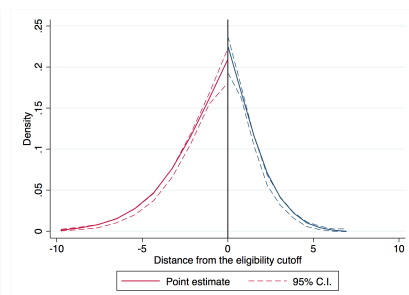
Figure A3: Manipulation testing



RV: n. of articles



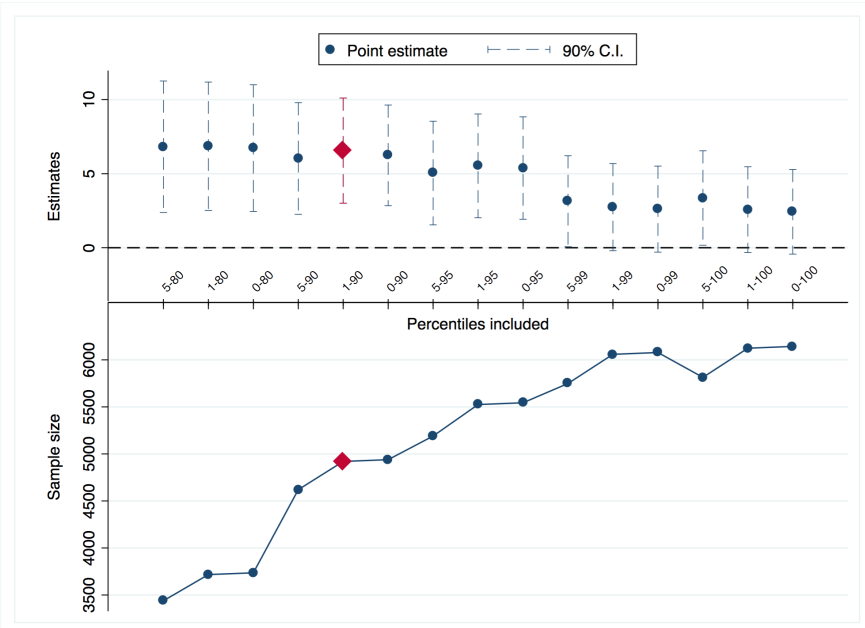
RV: n. of citations



RV: H-index

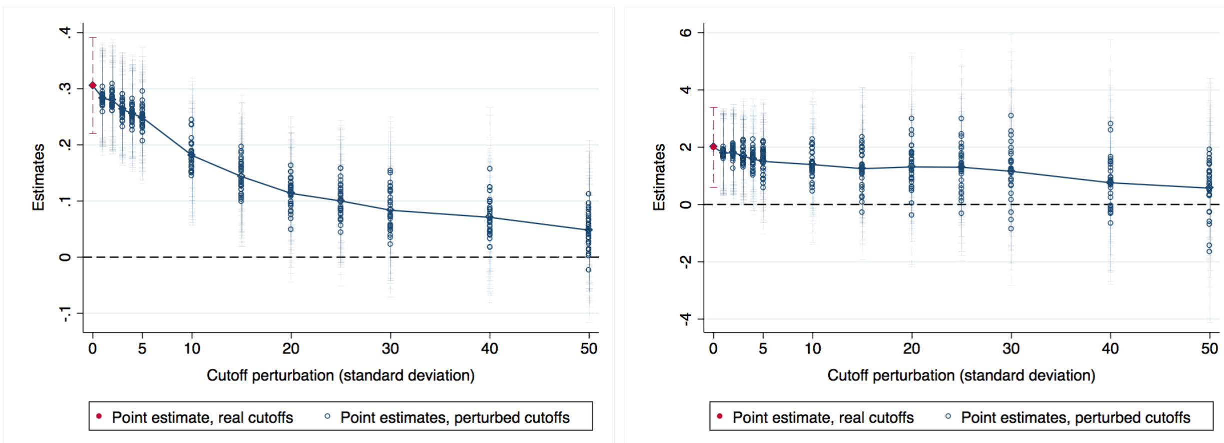
Notes. This figure depicts the local polynomial density estimation of candidates for the associate professor NSQ in 2012 in bibliometric fields, depending on their distance from each of the three bibliometric cutoffs. The local polynomial density is estimated following Cattaneo et al. (2017) and using the companion Stata package described in Cattaneo et al. (2018).

Figure A4: Robustness to sample restrictions



NOTES. This figure depicts, in the upper panel, the OLS coefficients of the indicator for the compliance with the two-out-of-three rule on the quantity of publications over the 2013-2016 period under different sample restrictions. The sample, before imposing the different restrictions, includes the candidates for the associate professor NSQ in 2012 in bibliometric fields. Each dot in the upper panel corresponds to the ITT coefficients estimated on specified inter-percentile ranges computed for candidates' distance from the relevant cutoffs. In each regression, the dependent variable is the total number of papers (including articles, conference papers, reviews and other items) published during the 2013-2016 period. The main independent variable is an indicator that equals one when a candidate complies with the two-out-of-three rule, that is, when her scores in at least two indicators are above the relevant cutoffs. The different sample restrictions are reported on the x-axis. For instance, when estimating the regression within the inter-percentile range '20-90', we exclude all candidates belonging to the top 10% or the bottom 20% of the pool of applicants in the same field for any of the three indicators considered. In the lower panel, the figures reports the sample size under the different sample restrictions. The red dots correspond to the sample chosen in our baseline specification. In each regression, within each academic field we exclude observations in the top 10% and the bottom 1% of the distribution of the distances. We also exclude the fields with more than 90% successful candidates and those with fewer than 30 observations. All regressions are estimated using a polynomial (quadratic) specification over the entire support.

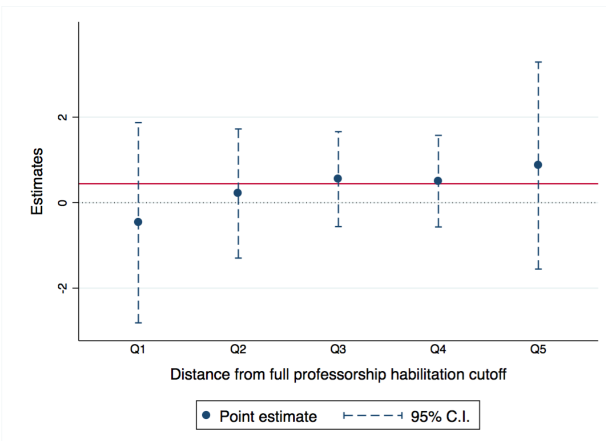
Figure A5: Placebo test: cutoff perturbation



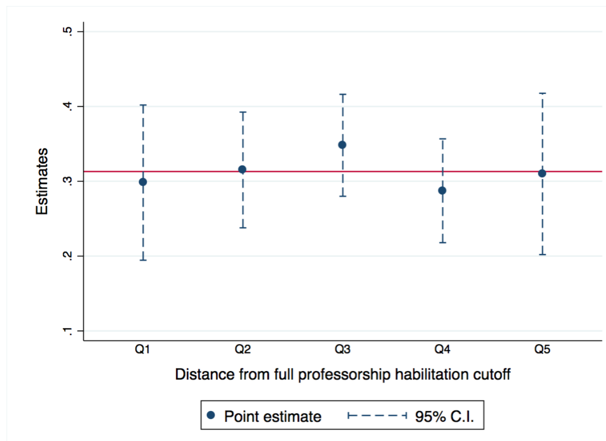
a. First stage **b.** Number of publications 2013-2016 (ITT)

NOTES. This figure depicts the OLS coefficients of the indicator for the compliance with the two-out-of-three rule on the outcome of the 2012 NSQ (Panel a) and on the quantity of publications over the 2013-2016 period (Panel B) applying different perturbations to the cutoff values. The sample includes the candidates for the associate professor NSQ in 2012 in bibliometric fields. Each dot corresponds to the coefficient of an indicator that equals one when a candidate complies with the two-out-of-three rule on the depended variable considered, under a different perturbation of the cutoff values. The dependent variables in Panels A and B are an indicator that equals one when a candidate obtains the qualification and the total number of papers (including articles, conference papers, reviews and other items) published during the 2013-2016 period, respectively. The permutations of the cutoff values are obtained by adding a randomly generated error component $\epsilon \sim N(0, \sigma)$, where the standard deviation (σ) of the error determines the intensity of the reshuffling. For each value of σ, we estimate 30 separate regressions for different realizations of ϵ. We apply the same reshuffling to the three bibliometric cutoffs, and we force the perturbation to lie within within - and +100% of the original cutoff values. In each regression, within each academic field, we exclude observations in the top 10% and the bottom 1% of the distribution of the distances. We also exclude the fields with more than 90% successful candidates and those with fewer than 30 observations. All regressions are estimated using a polynomial (quadratic) specification over the entire support.

Figure A6: Effect heterogeneity: distance from the full-professor thresholds



a. Number of publications (2009-2012)



b. Probability of achieving the associate professor NSQ (2012)

NOTES. This figure depicts the OLS coefficients of the interaction between the indicator for the compliance with the two-out-of-three rule and a set of indicators summarizing the heterogeneity in the distance from the full-professor thresholds on the quantity of publications over the 2013-2016 period. The sample includes the candidates for the associate professor NSQ in 2012 in bibliometric fields. The blue circles correspond to the coefficient of an interaction term formed by multiplying an indicator that equals one when a candidate complies with the two-out-of-three rule with a set of indicators that equal one if a candidate belongs to the specified quintile of the distribution of the distance from the full professor thresholds. The distance from the full professor threshold is expressed as a percentage of the initial stock of articles, thus defined as $dist_{i,1,s} = \frac{m_{1,s}^{full} - x_{i,1}}{x_{i,1}}$, where $m_{1,s}^{full}$ is the field-specific cutoff for the full professor NSQ and $x_{i,1}$ the professor's score in the same indicator. The dependent variable is the total number of papers (including articles, conference papers, reviews and other items) published during the 2013-2016 period. In each regression, within each academic field we exclude observations in the top 10% and the bottom 1% of the distribution of the distances. We also exclude the fields with more than 90% successful candidates and those with fewer than 30 observations. All regressions are estimated using a polynomial (quadratic) specification over the entire support.

A.2 Tables

Table A1: First stage - full professor NSQ

Candidates for full professor							
	Single RD (Articles)		Single RD (Citations)		Single RD (H-Index)		Triple RD
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Quadratic	LLR	Quadratic	LLR	Quadratic	LLR	Quadratic
Above cutoff	0.141*	0.175***	0.242***	0.174**	0.265***	0.312***	
	(0.072)	(0.064)	(0.075)	(0.081)	(0.086)	(0.108)	
Above 2/3 cutoffs							0.420***
							(0.067)
Academic field FE	Yes	No	Yes	No	Yes	No	Yes
Field specific interactions	Yes	No	Yes	No	Yes	No	Yes
Mean dep. var.	0.566	0.581	0.566	0.547	0.566	0.543	0.566
Sd dep. var.	0.496	0.494	0.496	0.498	0.496	0.498	0.496
BW (MSE)		0.298		0.270		0.182	
N of clusters	47	89	47	89	47	89	47
Observations	2369	1039	2369	675	2369	905	2369

Notes. This table reports the OLS coefficients of overcoming the field-specific bibliometric cutoffs on the outcome of the 2012 NSQ. The sample includes the candidates for the associate professor NSQ in 2012 in bibliometric fields. In all columns, the dependent variable is an indicator that equals one when a candidate gets the qualification, and zero otherwise. In Columns (1) to (4) the main independent variable is an indicator that equals one when a candidate overcomes the relevant cutoff for either the number of articles (Columns 1 and 2) or the number of citations (Columns 3 and 4). As for the h-index, estimates are not reported in this table. In Columns (5) and (6) the main independent variable is an indicator that equals one when a candidate complies with the two-out-of-three rule, that is, when her scores in at least two indicators are above the relevant cutoffs. Within each academic field, we exclude observations in the top 10% and the bottom 1% of the distribution of the distances. We also exclude the fields with more than 90% of successful candidatesand those with fewer than 30 observations. Regressions in Columns (1), (3), (5) and (6) are estimated using a polynomial (quadratic) specification over the entire support; in Columns (2) and (4) Local Linear Regressions (LLR) are estimated within the MSE-optimal bandwidth computed following Calonico et al. (2014), using the companion Stata package described in Calonico et al. (2017). In Column (6) Local Linear Regression is estimated within an arbitrary bandwidth, as the MSE-optimal bandwidth cannot be computed.
Standard errors, clustered at the academic field level, in parentheses. *** p< 0.01, ** p<0.05, and *p<0.10.

Table A2: Robustness check: local linear specification

Dependent variable: Number of publications		
	(1)	(2)
	Full Parametric	LLR
ITT	2.003***	1.559*
	(0.701)	(0.815)
LATE	6.557***	6.620*
	(2.159)	(3.718)
First Stage	0.306***	0.236***
	(0.043)	(0.056)
Academic field FE	Yes	Yes
Field specific interactions	Yes	No
Mean Dep. Var.	17.006	15.392
Standard dev.	20.410	11.950
BW (MSE) Ind 1		0.349
BW (MSE) Ind 2		0.465
BW (MSE) Ind 3		0.230
N. of clusters	82	84
Observations	4920	931

NOTES. This table reports the ITT and the LATE of achieving the qualification on the quantity of publications over the 2013-2016 period, and the corresponding first-stage estimates. In Column (1) we use our baseline full-parametric specification over the entire support. In Column (2), we use a linear approach in the neighborhood of the threshold. The sample includes the candidates for the associate professor NSQ in 2012 in bibliometric fields. The dependent variable in both columns is the total number of papers (including articles, conference papers, reviews and other items) published during the 2013-2016 period. The ITT is the OLS coefficient of the indicator that equals one when a candidate complies with the two-out-of-three rule; the LATE is the 2SLS coefficient of the indicator that equals one when a candidate obtains the qualification. The latter is instrumented by the indicator that equals one when a candidate complies with the two-out-of-three rule. The polynomial (quadratic) specification in Column (1) is estimated over the entire support after excluding observations in the top 10% and the bottom 1% of the distribution of the distances. The local linear specification in Column (2) is estimated within a joint three-dimensional bandwidth. The bandwidth for each productivity indicator is the MSE-optimal bandwidth computed following Calonico et al. (2014) and using the companion Stata package described in Calonico et al. (2017). In both columns, we exclude the fields with more than 90% successful candidates and those with fewer than 30 observations.
Standard errors, clustered at the academic field level, in parentheses. *** p< 0.01, ** p<0.05, and *p<0.10.

Table A3: Robustness check - candidates applying to the competition sector they belong to as of 2012.

Candidates for associate professor				
	Publications	Articles	Conf. Papers	Reviews
	(1)	(2)	(3)	(4)
ITT	1.520**	1.230**	0.218	0.045
	(0.761)	(0.533)	(0.290)	(0.109)
LATE	4.238**	3.431**	0.609	0.127
	(1.951)	(1.376)	(0.709)	(0.263)
Mean Dep. Var.	15.374	11.285	2.216	0.895
Standard dev.	13.728	10.111	5.719	2.048
N. of clusters	83	83	83	83
Observations	5024	5024	5024	5024

NOTES. This table reports the Intention-to-treat (ITT) and the Local Average Treatment Effect (LATE) of achieving the qualification on the quantity of publications over the period 2013-2016. The sample includes the candidates for the associate professor NSQ in 2012 in bibliometric fields. In case of multiple applications, we consider the one to the competition sector to which the scholar already belongs as an assistant professor as of December 2012. The dependent variable in Column (1) is the total number of papers (including articles, conference papers, reviews and other items) published during the period 2013-2016; the dependent variables in Columns (2), (3) and (4) are the total number of articles, conference papers and reviews published during the period 2013-2016, respectively. The ITT is the OLS coefficient of the indicator that equals one when a candidate complies with the two-out-of-three rule; the LATE is the 2SLS coefficient of the indicator that equals one when a candidate gets the qualification. The latter is instrumented by the indicator that equals one when a candidate complies with the two-out-of-three rule. Within each academic field, we exclude observations in the top 10% and the bottom 1% of the distribution of the distances. We also exclude the fields with more than 90% successful candidates and those with fewer than 30 observations. All regressions are estimated using a polynomial (quadratic) specification over the entire support.
Standard errors, clustered at the academic field level, in parentheses. *** $p< 0.01$, ** $p<0.05$, and *$p<0.10$.

Table A4: The effect of achieving the NSQ on the number of publications, per year

Panel A - Dependent variable: Number of publications								
	2009	2010	2011	2012	2013	2014	2015	2016
ITT	0.149	0.138	0.008	0.253	0.474**	0.496**	0.323	0.717***
	(0.154)	(0.167)	(0.153)	(0.222)	(0.236)	(0.211)	(0.273)	(0.254)
LATE	0.486	0.452	0.025	0.827	1.556**	1.626***	1.058	2.348***
	(0.446)	(0.482)	(0.430)	(0.634)	(0.679)	(0.618)	(0.801)	(0.783)
Mean Dep. Var.	2.940	2.962	3.146	3.533	3.820	4.159	4.518	4.574
Standard dev.	3.015	3.019	3.141	3.793	3.862	4.956	7.547	8.446
Observations	4827	4843	4855	4861	4869	4898	4914	4919

Panel B - Dependent variable: No publications								
	2009	2010	2011	2012	2013	2014	2015	2016
ITT	-0.043	-0.029	0.011	-0.029	-0.050	-0.019	-0.005	-0.032
	(0.027)	(0.031)	(0.029)	(0.022)	(0.030)	(0.023)	(0.028)	(0.026)
LATE	-0.140*	-0.094	0.036	-0.096	-0.162*	-0.063	-0.015	-0.103
	(0.078)	(0.092)	(0.079)	(0.063)	(0.086)	(0.066)	(0.079)	(0.074)
Mean Dep. Var.	0.147	0.160	0.153	0.125	0.136	0.123	0.128	0.129
Standard dev.	0.354	0.366	0.360	0.331	0.343	0.328	0.334	0.336
Observations	4920	4920	4920	4920	4920	4920	4920	4920

Panel C - Dependent variable: Number of articles								
	2009	2010	2011	2012	2013	2014	2015	2016
ITT	0.056	0.097	0.079	0.169	0.231	0.342*	0.160	0.530***
	(0.120)	(0.117)	(0.114)	(0.165)	(0.161)	(0.174)	(0.204)	(0.186)
LATE	0.183	0.316	0.258	0.555	0.758	1.121**	0.526	1.735***
	(0.338)	(0.336)	(0.326)	(0.474)	(0.468)	(0.514)	(0.590)	(0.567)
Mean Dep. Var.	2.084	2.041	2.206	2.468	2.734	3.054	3.408	3.598
Standard dev.	2.339	2.229	2.327	2.979	2.849	4.183	6.854	7.807
Observations	4827	4843	4855	4861	4869	4898	4914	4919

Notes. This table reports the ITT and the LATE of achieving the NSQ on the quantity of publications in each year over the 2009-2016 period. The sample includes the candidates for the associate professor NSQ in 2012 in bibliometric fields. The dependent variables in Panels A and C are the numbers of publications and articles, respectively. The dependent variable in Panel B is an indicator that equals one when a candidate does not publish any paper in a given year. The ITT is the OLS coefficient of the indicator that equals one when a candidate complies with the two-out-of-three rule; the LATE is the 2SLS coefficient of the indicator that equals one when a candidate achieves the NSQ. The latter is instrumented by the indicator that equals one when a candidate complies with the two-out-of-three rule. First-stage estimates are reported in Column (7) of Table III. In Panels A and C, the sample includes only scholars that were 'active' in the year considered, that is, scholars whose first publication is not later than that year. Within each academic field, we exclude observations in the top 10% and the bottom 1% of the distribution of the distances. We also exclude the fields with more than 90% successful candidates and those with fewer than 30 observations. All regressions are estimated using a polynomial (quadratic) specification over the entire support. Standard errors, clustered at the academic field level (N=82), in parentheses. *** p< 0.01, ** p<0.05, and *p<0.10.

Table A5: Effect heterogeneity: main subject

Dependent variable: Number of publications										
	MATH	PHYS	CHEM	EARTH	BIOL	MED	AGRO	ENG	ARCH	PSY
Panel A: LATE										
qualified × subject	1.338	50.127***	13.437*	8.069**	0.082	3.582	19.019	20.410*	19.064	-3.047
	(2.764)	(8.913)	(7.451)	(3.865)	(1.317)	(2.476)	(18.982)	(11.371)	(30.293)	(4.808)
Mean (subject)	11.646	39.790	17.721	13.601	11.638	16.691	12.519	14.379	21.150	11.897
St. dev. (subject)	11.014	56.737	10.948	7.949	8.194	15.777	9.470	10.079	15.106	10.291
Panel B: First Stage										
above 2/3 cutoffs × subject	0.372***	0.143***	0.242***	0.507**	0.313***	0.394***	0.186*	0.252	0.114	0.532**
	(0.109)	(0.051)	(0.064)	(0.243)	(0.074)	(0.096)	(0.105)	(0.206)	(0.140)	(0.215)
Mean (subject)	0.553	0.790	0.703	0.710	0.523	0.445	0.672	0.551	0.703	0.623
St. dev. (subject)	0.498	0.408	0.458	0.455	0.500	0.497	0.470	0.499	0.458	0.486

NOTES. This table reports the OLS coefficients of the indicator for compliance with the two-out-of-three rule and its interaction with a set of subject-specific dummies on the quantity of publications over the 2013-2016 period (Panel A) and on the outcome of the 2012 NSQ (Panel B). The sample includes the candidates for the associate professor NSQ in 2012 in bibliometric fields. The dependent variable in Panel A is the total number of papers (including articles, conference papers, reviews and other items) published during the 2013-2016 period; the dependent variable in Panel B is an indicator that equals one when a candidate obtains the qualification, and zero otherwise. In both panels, the main independent variables are a set of interaction terms formed by multiplying an indicator that equals one when a candidate complies with the two-out-of-three rule with each of the subject-specific dummies, that is, a set of indicators equal to 1 if a candidate belongs to the specified subject. Within each academic field, we exclude observations in the top 10% and the bottom 1% of the distribution of the distances. We also exclude the fields with more than 90% successful candidates and those with fewer than 30 observations. All regressions are estimated using a polynomial (quadratic) specification over the entire support.
Standard errors, clustered at the academic field level, in parentheses. *** $p< 0.01$, ** $p<0.05$, and *$p<0.10$.
Legend: MATH=Mathematics; PHYS=Physics; CHEM=Chemistry; EARTH=Earth Sciences; BIOL= Biology; MED=Health Sciences; AGRO=Agronomy and Veterinary; ENG=Engineering; ARCH=Architecture; PSY=Psychology.