The Italian Reform of the academic recruitment system: an appraisal of ANVUR and CUN benchmarks for assessing candidates and commissioners

Abatemarco, Antonio and Dell'Anno, Roberto

2012

The Italian Reform of the academic recruitment system.

An appraisal of ANVUR and CUN benchmarks for assessing candidates and commissioners

Antonio Abatemarco[a] **Roberto Dell'Anno**[b]

[a] Università di Salerno - *Dipartimento di Economia e Statistica -Via Ponte Don Melillo, 1- 84084 Fisciano (Salerno).*

[b] Università di Foggia - *Dipartimento di Scienze Economiche, Statistiche e Matematiche –Largo Papa Giovanni Paolo II, 1- 71100 - Foggia.* **(autore corrispondente)**

Abstract

The present study is aimed at contributing to the ongoing debate about the implications of the incoming recruitment system as proposed by Law 240/2010 (Gelmini's reform). For this purpose, the main implications of the two alternative criteria respectively proposed by the National Agency for the Evaluation of the University System and Research (ANVUR) and the National University Council (Consiglio Universitario Nazionale - CUN) are investigated for assessing candidates and commissioners admitted to apply for the national scientific approval of the Italian academic recruitment system.
Using the sample of 1327 Italian academic economists (secs p/01, p/02, p/03) enrolled for the academic year 2011-12, the analysis provides two simulations. First, the thresholds for both ANVUR and CUN criteria as well as the resulting shares of "qualified" candidates and commissioners are computed. Second, the impact of the new eligibility criteria on the academic competitions (p01, p02, p03) that occurred in 2005 is simulated under the assumption that behavioral responses are absent.
The findings suggest that (i) CUN criteria provide more selective benchmarks than ANVUR ones, and (ii) in the absence of behavioral responses, the new system of recruitment is expected to remarkably affect the profile of the Italian academic system.

JEL codes: *I28, O38*

Keywords: *Academic recruitment system; research assessment; citation-based indices; Italian University Reform*

Acknowledgment: This paper has benefited from the comments and suggestions of the editor and two anonymous referees. The usual disclaimer applies. The paper was partly written when R. Dell'Anno was visiting at the Andrew Young School of Policy Studies at Georgia State University. The hospitality of this Institution is gratefully acknowledged.

1. Introduction

The debate on the recruitment system is a key issue for the academic system, as the quality of human capital is definitely the most relevant factor of production for these organizations. In addition, it is clear that the definition of recruitment criteria is important for other general economic subjects such as the allocation of public resources, the creation of human capital in the economic system, the process of knowledge creation in terms of (dis)incentives to scientific investigation on some themes instead of others, and so on.[1]

According to OECD (2011), there is a major flaw in the Italian teaching staff management system with regard to the recruitment procedures. A vast literature provides anecdotal evidences on the lack of meritocracy and nepotism that might have distorted the results of some academic recruitment procedures.[2] In this scenario, several reforms have been approved in the last decade for improving the academic recruitment mainly through the reduction of space/incentives for manipulation and pressure. For a better understanding of the intrinsic motivations underlying the upcoming recruitment procedure (L. 240/2010, also called *riforma Gelmini*), it is worth providing a brief description of the evolution of the Italian academic system in the last decade.

Basically, until 1998 career progressions were made in a nationwide public competition where (i) the panel of commissioners was elected by the whole body of professors of that academic sub-discipline (hereinafter referred to as *SSD*[3]), (ii) the panel identified the "qualified" candidates (*idonei*), and (iii) each University with a vacancy was entitled to hire (*chiamare*) by picking from the lists of qualified candidates.[4]

Law 210/1998 brought about an in-depth reformation of the recruitment system by moving towards a decentralized mechanism. This reform was inspired by the belief that decentralization of the recruitment strategies would have made Universities responsible and, therefore, more prone to reward meritocracy. According to this system, career progressions were made in local public competitions so that each University was entitled to manage its own *concorso* to fill its vacancies. Each commission could declare up to 3 (later reduced to 2 and then 1) *idonei* to be eventually *hired* by any Italian University. Unfortunately, replacing the centralized system with a decentralized one further increased the lack of meritocracy. Almost immediately after the 1998 reform, part of the academic staff and public opinion realized that this system was not sustainable. This belief motivated several legislative attempts (e.g.

---

[1] The debate on academic recruitment assumes a greater role in these years, mainly as a consequence of both un-balanced recruitment policies followed by the academic system in the previous decades and the relevant reduction of associate and full professors started since 2007. In particular, according to CNVSU (2010), until 2015 about 19% of Full professors and 5% of associate and assistant professors would retire at retirement age. Taking into account that this kind of retirement is weighted towards about 50% of the effective dismissed professors, thus, the new recruitment system will affect strongly the future of the Italian Universities.

[2] See, for an overview, Perotti (2002a, 2008), Durante et al. (2011) and OECD (2011).

[3] It stands for *Settore Scientifico Disciplinare* (Scientific Disciplinary Sector).

[4] See Checchi (1999) for an appraisal of this system based on data for associate professorship that took place in Italy during the academic year 1997-98.

D.P.R. 117/2000; L. 230/2005; D.Lgs. 164/2006) to improve the recruitment system; however, each time it resulted in unsatisfactory results.

According to OECD (2011), centralized and decentralized recruiting procedures have shown their pitfalls and limitations and none have succeeded in ensuring a consistent selection of the "best" candidates. Centralized recruiting procedures mainly have the virtue of weakening local interference and manipulating in the decision process. However, the national panel of experts has little incentive to declare only the "best" candidates, as it does not have any strong direct interests to do so. On the other hand, a decentralised system has also failed on several aspects. For the economic area, Perotti (2002b, 2003, 2004, 2006) shows that these recruitment procedures have been often characterized by the lack of meritocracy and manipulation. It is well known that local competitions have provided a disproportionate advantage to "home" candidates, thereby discouraging applications from larger pools of qualified scholars. It has been documented in official statistics that the large majority of successful candidates to academic positions are internal. For instance, internal *idonei* over the period 1999-2010 are, on average, 81.5% and 77.1% for full and associate professors, respectively[5] (CNVSU - National Committee for the Assessment of the University System, 2010). Gagliarducci et al. (2005), Perotti (2008), and OECD (2011) have indicated failure of the Italian academic recruitment system in the lacking of incentives/penalties for the examiners/Universities to select according to meritocracy. They have suggested breaking off the decline of Italian Universities to move toward performance-based careers and salary progression.[6]

Law 240/2010 deeply reforms the academic recruitment system by introducing a two-step procedure. The first step occurs at the national level. It is aimed at forming a list of national scientific qualified scholars (for each SSD) eligible to be hired as associate or full professor. This *"abilitazione scientifica"* would be granted, yearly, by a national panel composed of five members, four of whom will be drawn from lists of full professors grouped by discipline (SSD) and one drawn from a list of foreign scholars and experts proposed by the National Agency for the Evaluation of the University System and Research (ANVUR). This step of the recruitment process would open every October for a period of five months and the resulting "qualified" (*abilitati*) will be valid for no more than four years.

The second step of recruitment occurs at the decentralized level. Each University (i.e. Department) decides whether and how to fill its vacancies by advertising professorship positions. The sole qualified scholars, or professors with an equivalent position in foreign Universities, are entitled to apply. Evidently, this step of the procedure ruled by L. 240/2010 is crucial as well: the decision at the local level about

[5] Perotti (2008) analyzes 40 academic competitions to full professorship in Economics, between 1999 and 2002. In 32 of these competitions, at least one winner was an internal; in total, 35 internals were declared winners. Of these, 27 were appointed full professors in their own University and only 5 in a different University. Durante et al. (2011) estimate the probability to win, conditioned to the status of "home" or "visitors" candidate. The results are incontrovertible: if you are an internal candidate your probability to win is 0.57 against only 0.27 if you are employed outside the University that issued the call to fill the vacancy.

[6] Academic salaries are uniform nationally for each level of professorship and these are largely determined by seniority. On these aspects, minor changes have been introduced by L. 240/2010.

which vacancy to advertise (among different SSDs) and which *abilitato* to hire (within the same SSD) may be, once again, distorted by opportunistic behaviours.[7]

This paper focuses on the most innovative mechanism proposed by this reform, that is, the setting up of a set of bibliometric benchmarks to pick candidates. The system has the following two aims: (a) to limit the size of applicants admitted for peer-review evaluation to the scholars with adequate publication records and, (b) to reduce the discretionary power of a national panel of experts. In comparison with the previous systems, any opportunistic or nepotistic behaviour can be applied only to scholars holding an adequate scientific production.[8]

At present, two alternative benchmark systems have been proposed for selecting the candidates, respectively, by the ANVUR and the National University Council (*Consiglio Universitario Nazionale* - CUN). The criteria and benchmarks differ from each other among scientific macro-areas and, sometimes, among SSDs as well (e.g., Area 13 "Economics and Statistics"- CUN's classification). For the purposes of this study, we focus on micro, macro, and public economists corresponding to SSDs, that is, secs-p/01, secs-p/02, and secs-p/03, respectively.[9]

The present article is aimed at contributing to the ongoing debate on the recruitment procedure from a double perspective. Firstly, we estimate the implications of the incoming recruitment system as proposed by L. 240/2010. Secondly, we attempt to evaluate the ways in which the incoming recruitment system could effectively change the criteria to recruit professors in the following years.

The paper is organized as follows. Section 2 describes the main characteristics of bibliometric archives used in this work with special reference to the economic area. Sections 3 and 4 explain the benchmark criteria proposed by ANVUR and CUN for selecting candidates for peer-review evaluation and commissioners to be eligible for a national panel, respectively. Section 5 compares the results of 49 competitions advertised in 2005 with simulated findings based on ANVUR criteria. In addition, we analyse the post-*concorso* scientific performances of winners and losers as of December 2011. Section 6 provides some hints to potentially improve ANVUR criteria for selecting candidates and commissioners. Finally, section 7 provides the conclusion.

2. The database

In the past decades there has been an increased attraction of bibliometric indexes for evaluating the quantity, quality, and impact of scientific productivity. These indexes do not account for the "internal" quality of scientific product (e.g. article, books) or the ability of producer (e.g. researcher, department);

[7] Here, the more accountable and responsible are the University/department for the performances of hired professors, the larger the incentive to follow meritocracy by choosing among *abilitati*.

[8] Moreover, in order to reduce manipulation, each panel is also constrained to evaluate eligible candidates according to a set of criteria defined by ANVUR and made publicly available in advance.

[9] Also, Perotti focuses on these three SSDs to analyse the academic recruitment system for Italian economists.

however, they provide useful indications about the degree of "utilization" (impact) that these objects/subjects have received by peers in their area of research.

The success of the bibliometric approach to scientific evaluation has increased mainly owing to its greater availability and the exhaustiveness of citation archives. Evidently, the reliability and robustness of the citation index is increasing with the quality and quantity of data it utilizes. In this sense, the choice of the database from which to extract bibliometric data plays a crucial role in the interpretation of results and the reliability of any ranking arising from bibliometric indicators.

In general, we distinguish two categories of citation databases: (a) databases that include only products that have passed scientific peer-review, and (b) archives that also include products without any sort of qualitative assessment of peers.

In the first category there are two internationally recognized databases, the Institute for Scientific Information-Web of Science owned by Thomson Reuters[10] (hereinafter, WoS), and, since 2004, Scopus published by the international publisher Elsevier.[11]

In the second category, with reference to Economics, the major archives are Google Scholar, EconLit, and RePEc. The first one, owned by Google Incorporation, indexes articles, theses, books, abstracts and court opinions, from academic publishers, professional societies, online repositories, Universities, and other web sites.[12] It offers the widest coverage of academic communication; however, it can be strongly manipulated through self-citations. Further source of problems for Google Scholar is that it is not built from structured records provided by publishers (metadata) but it analyzes full text of the publications (Hicks and Wang, 2009).[13]

EconLit is published by the American Economic Association, and it provides a bibliographic coverage of a wide range of economics-related literature. It has access to over 800,000 records. Six types of documents (written not only in English) are indexed in EconLit: journal articles, books, collective volume articles, dissertations, working papers in economics, and book reviews from the Journal of Economic Literature.[14]

RePEc (*Research Paper in Economics*) accesses over 1,000,000 items. It is a decentralized database of working papers, journal articles, books, and software components.[15]

[10] http://science.thomsonreuters.com/
[11] http://www.scopus.com/home.url
[12] http://scholar.google.com/intl/en/scholar/about.html
[13] For instance, Google Scholar ignores reference lists if the keywords "Bibliography" or "References" is absent in the document (Hicks and Wang 2009, p. 9).
[14] http://www.aeaweb.org/econlit/index.php
[15] http://repec.org/

In this research three bibliometric archives will be used: Scopus and WoS, as they are expressly indicated as official data sources by ANVUR; and EconLit (search limited to journal articles), as it is the widest archive of peer-review journal articles in Economics.

According to the authors, EconLit is a better choice in comparison to the other archives preferred when benchmarks are not related to citational profiles of scholars (e.g. CUN's criteria).[16]

The comparison of citation databases has been examined by several studies. In general, it has been shown that the coverage of Scopus and WoS varies considerably between scientific disciplines (Meho and Yang, 2007; Harzing, 2008, 2010; Bar -Ilan, 2008).

In addition to the number of indexed journals, a very relevant factor is the temporal coverage of the citations. On this issue, the WoS citation indexes since 1990, Scopus indexes from 1996, while EconLit does not provide citation statistics.[17]

To our knowledge there is no literature that has attempted to evaluate the bibliometric database having the widest coverage for Italian economists. Our empirical analysis allows evaluating both the differences and their impact on the recruitment procedures.

In summary, there is strong empirical evidence that EconLit has the widest coverage of scientific production for the Italian academic economists. Looking at the citation archives, Scopus accounts considerably more for both, documents and citations than WoS in Economics.

In particular, with the reference to 1,327 scholars analyzed between November and December 2011: EconLit indexed 12,022 publications and had no records for 123 economists (9.3% of the sample); Scopus[18] accounted for 5,951 publications (from 2002 to 2011), 49,847 citations, and had no records for 351 economists (26% of the sample); and WoS[19] indexed 3,803 publications (from 2002 to 2011), 36,107 citations, and had no records for 411 professors (31% of the sample).[20]

[16] In absolute terms, Google scholar is the widest database but it has two main drawbacks. It also includes non peer-reviewed documents and it is quite easily prone to manipulation (see on this issue Labbé 2010), which implies that Google scholar can be used only jointly with other peer-review archives. These reasons are likely to motivate ANVUR's (2011b) choice to not expressly cite Google scholar among the data sources for the bibliometric assessment of the candidates.

[17] Tarantino (2006) points out that, although Scopus indexes publications subsequent to its main competitor (WoS), its search engine is able to track up to items published in the Sixties when cited after 1995.

[18] Bibliometric data are extracted by limiting the fields to Social Sciences and Humanities. These areas include: arts and humanities; business, management and accounting; decisions sciences; economics, econometrics and finance, psychology, social sciences, multidisciplinary.

[19] Bibliometric data are extracted by limiting the search to the following scientific areas: economics, industrial relations labour, political science, statistic probability, social science mathematical methods, business finance, history of social science, mathematics interdisciplinary, planning, development applications, social issues, ethics, international relations, social sciences interdisciplinary, environmental studies.

[20] As data are extracted at individual level these values contain duplications as consequence of co-authorships.

3. ANVUR and CUN eligibility criteria for associate and full professorships

The 2010 University Reform Act (Law 240/2010) is aimed at improving the meritocracy within the academic recruitment system. It introduces a two-step procedure for the selection of associate and full professors. It partially moves from a system of competitive local vacancies into a national certification system. The scientific qualification "*abilitazione*" is declared by a national commission; however, the selection of the qualified scholar to be engaged is discretionally made by each local institution (i.e. Department/University).

This paragraph focuses on the set of bibliometric benchmarks to select among candidates who are eligible for the peer-review evaluation of the national commission.

We analyze how the two main alternative benchmarks proposed by the ANVUR and CUN might affect the recruitment procedures.[21]

In addition, the differences between ANVUR and CUN criteria reveal a theoretical-based position on the reliability of citation metrics as evaluation criteria to assess research activity. This is strictly related with an epistemological issue that considers Economics as a discipline on the borderline between "strong" sciences and "weak" (social) sciences.

Specifically, although both organisms suggest quite different benchmarks for each scientific area, CUN (2011) excludes the citation indexes among the criteria for the economic assessment of the candidates, while ANVUR asserts their use for some scientific disciplines as a subset of economic disciplines like secs-p/01, secs-p/02, and secs-p/03.

In the following two sub-sections we provide details on the two benchmark systems and comment on their implications.

3.1 ANVUR eligibility criteria for associate and full professorships

ANVUR (2011b) states that in order to apply for national peer-review evaluation candidates must hold both the requisites (i) to improve (median) quality of Italian scientific research, and (ii) to show a reasonable continuity of their publication records.

The first necessary requisite is related to the "quality" of scientific production. It requires that the candidates for career progression must hold at least two of the three specified bibliometric indexes higher than the median calculated on the samples of associate or full professors of their respective SSD. This "quantity/impact" constraint is held if at least two of the following benchmarks are satisfied by the candidate:

[21] At the time when this paper was being written (March 2011) the Ministry for the University and Research (MIUR) had not yet opted for the method that will be applied. When the article has been accepted for publication the policy maker has opted for benchmarks proposed by the ANVUR.

- 1° criterion: the number of journal articles and monographs indexed by WoS or Scopus, or any other database that has characteristics of wide coverage and methodological rigor, in the last ten years. This parameter should be appropriately normalized to the academic age if it is less than ten years;[22]
- 2° criterion: the total number of citations;
- 3° criterion: the h-index (Hirsch, 2005). This index may be supported or replaced with other indexes under evaluation by ANVUR (e.g. h-IF).

Concerning the 1° criterion, ANVUR does not explain what kind of normalization should be applied. We fill this gap by applying, if the year of the first indexed publication ($year_0$) is prior to 2002, the following formula: $n\left[10/\left(2012 - year_0\right)\right]$, where, n is the number of publications.

Tables 1 and 2 present the median values of three criteria for each SSD estimated by Scopus and WoS databases, respectively.

Table 1: Medians according to ANVUR criteria estimated using Scopus database (2011) - candidates

	Assistant Professors			Associate Professors			Full Professors		
	1°crit. (# publ.)	**2°crit. (# cit.)**	**3°crit. (h-)**	**1°crit. (# publ.)**	**2°crit. (# cit.)**	**3°crit. (h-)**	**1°crit. (# publ.)**	**2°crit. (# cit.)**	**3°crit. (h-)**
p/01	5.0	1.0	1.0	5.0	8.0	2.0	5.0	5.0	2.0
p/02	3.3	0.0	0.5	4.0	8.0	1.0	4.0	4.0	1.0
p/03	3.3	0.0	0.0	4.0	3.0	1.0	1.0	1.0	1.0

Table 2: Medians according to ANVUR criteria estimated using WoS database (2011) - candidates

	Assistant Professors			Associate Professors			Full Professors		
	1°crit. (# publ.)	**2°crit. (# cit.)**	**3°crit. (h-)**	**1°crit. (# publ.)**	**2°crit. (# cit.)**	**3°crit. (h-)**	**1°crit. (# publ.)**	**2°crit. (# cit.)**	**3°crit. (h-)**
p/01	3.3	0.0	0.0	3.3	4.0	1.0	3.3	3.3	2.0
p/02	1.5	0.0	0.0	2.0	3.0	1.0	1.7	1.7	1.0
p/03	3.3	0.0	0.0	2.0	3.0	1.0	0.0	0.0	1.0

It can be pointed out from Tables 1 and 2 that:

(a) The median values are generally higher for associates than for full professors.

(b) By comparing the medians for full professors among SSDs, there is evidence that p/01 has the highest medians, while p/03 has the lowest. These differences are larger for full professors.

(c) Scopus database has wider coverage, and therefore, higher bibliometric scores than WoS.

The second necessary requirement (hereinafter, 4th criterion) asks for "reasonable" continuity in scientific production. It implies that to be admitted for peer-review evaluation by a national panel, applicants must hold a "reasonable" continuity in scientific production measured over the last 5 years to become associate

[22] ANVUR defines the first year of academic age as the year of the first publication indexed by WoS or Scopus. ANVUR (2011b) motivates this choice by operative reasons. In particular, the lack of data about more reliable information on academic age (e.g. the year of PhD attainment).

professor and in the last 10 years to become full professor. However, ANVUR (2011a, b) does not define quantitatively the meaning of "reasonable" continuity. This indefiniteness gives flexibility to the first step of recruitment process. For instance, if the share of eligible candidates is excessively high (or low) within SSD or excessively different compared to similar SSDs, thus policy maker has the chance to fine tuning the shares of candidates admitted to national peer-review evaluation. Table 3 tests the sensitiveness of the eligibility criteria by changing the constraints on continuity of scientific production.

Table 3: Eligible candidates according to alternative benchmarks for the 4th criterion (ANVUR)

	Benchmarks		**Scopus**		**WoS**	
To:	**Associate**	**Full**	**Associate**	**Full**	**Associate**	**Full**
From:	(5 years)	(10 years)	(total)	(total)	(total)	(total)
Assistant	**# publ.≥ 1**	**# publ. ≥ 2**	116	146	113	127
Associate		**# publ. ≥ 2**	-	162	-	138
Assistant	**# publ.≥ 2**	**# publ.≥ 4**	111	124	97	78
Associate		**# publ.≥ 4**	-	143	-	107
Assistant	**# publ.≥ 3**	**# publ.≥ 6**	104	90	71	34
Associate		**# publ.≥ 6**	-	118	-	61
Assistant	**# publ.≥ 4**	**# publ.≥ 8**	94	54	60	13
Associate		**# publ.≥ 8**	-	77	-	35
Assistant	**# publ.≥ 5**	**# publ.≥10**	81	30	38	5
Associate		**# publ.≥10**	-	53	-	22

In our view, an excessive selection based on the "continuity" criterion might be distortive as it encourages the quantity, instead of the quality, of scientific research. Moreover, especially for full professorship, the fourth criterion is based on the same information used for the first one and therefore the number of recent publications would assume a key role in the skimming process. Based on this view, we assume that to apply for associate (full) professorship, the candidates must have at least two (four) international publications indexed by Scopus or WoS over the period 2007-2011 (2002-2011). These values are chosen by looking at the benchmarks proposed by CUN.

Table 4 reports the number and the percentages in their respective SSD and academic roles of applicants who will be admitted to peer-review evaluation of the national panel on the basis of Scopus archive.

Table 4: Eligible candidates according to ANVUR criteria (estimated using Scopus database)

To:	Associate Professor					Full Professor				
From:	**1°crit.**	**2°crit.**	**3°crit.**	**4°crit.**	**Eligible**	**1°crit.**	**2°crit.**	**3°crit.**	**4°crit.**	**Eligible**
p/01										
Assistant Prof. (310)	144 (46%)	73 (24%)	30 (10%)	168 (54%)	**61 (20%)**	144 (46%)	102 (33%)	30 (10%)	113 (36%)	**76 (25%)**
Associate Prof. (211)	-	-	-	-	**-**	94 (45%)	124 (59%)	66 (31%)	106 (50%)	**88 (42%)**
p/02										
Assistant Prof. (128)	58 (45%)	30 (23%)	36 (28%)	59 (46%)	**36 (28%)**	58 (45%)	43 (34%)	36 (28%)	44 (34%)	**33 (26%)**
Associate Prof. (95)	-	-	-	-	**-**	47 (49%)	57 (60%)	36 (49%)	46 (48%)	**43 (45%)**
p/03										
Assistant Prof. (69)	32 (46%)	17 (25%)	14 (20%)	29 (42%)	**14 (20%)**	42 (61%)	22 (32%)	14 (20%)	16 (23%)	**15 (22%)**
Associate Prof. (37)	-	-	-	-	**-**	25 (68%)	23 (62%)	16 (43%)	13 (35%)	**12 (32%)**
Total										
Assistant Prof. (507)	234 (46%)	120 (24%)	80 (16%)	256 (50%)	**111 (22%)**	244 (48%)	167 (33%)	80 (16%)	173 (34%)	**124 (24%)**
Associate Prof. (343)	-	-	-	-	**-**	166 (48%)	204 (59%)	129 (38%)	165 (48%)	**143 (42%)**

Table 5 presents the number and the percentages in their respective SSD and academic roles of applicants who will be admitted to peer-review evaluation of the national panel on the basis of the WoS database.

Table 5: Eligible candidates according to ANVUR criteria (estimated using WoS database)

To:	Associate Professor					Full Professor				
From:	**1°crit.**	**2°crit.**	**3°crit.**	**4°crit.**	**Eligible**	**1°crit.**	**2°crit.**	**3°crit.**	**4°crit.**	**Eligible**
p/01										
Assistant Prof. (310)	146 (47%)	71 (23%)	54 (17%)	127 (41%)	**65 (21%)**	146 (47%)	84 (27%)	7 (2%)	67 (22%)	**50 (16%)**
Associate Prof. (211)	-	-	-	-	**-**	104 (49%)	106 (50%)	34 (16%)	78 (37%)	**70 (33%)**
p/02										
Assistant Prof. (128)	60 (47%)	30 (23%)	22 (17%)	39 (30%)	**23 (18%)**	62 (48%)	39 (30%)	22 (17%)	20 (16%)	**17 (13%)**
Associate Prof. (95)	-	-	-	-	**-**	54 (57%)	57 (60%)	32 (34%)	28 (29%)	**28 (29%)**
p/03										
Assistant Prof. (69)	38 (55%)	12 (17%)	8 (12%)	22 (32%)	**9 (13%)**	40 (58%)	25 (36%)	8 (12%)	11 (16%)	**11 (16%)**
Associate Prof. (37)	-	-	-	-	**-**	26 (70%)	24 (65%)	12 (32%)	9 (24%)	**9 (24%)**
Total										
Assistant Prof. (507)	244 (48%)	113 (22%)	84 (17%)	188 (37%)	**97 (19%)**	248 (49%)	148 (29%)	37 (7%)	98 (19%)	**78 (15%)**
Associate Prof. (343)	-	-	-	-	**-**	184 (54%)	187 (55%)	78 (23%)	115 (34%)	**107 (31%)**

Tables 4 and 5 provide empirical evidence for the following:

(a) On aggregate, the choice of bibliometric databases (Scopus or WoS) does not significantly affect number of selected candidates. It is consistent to the ANVUR choice to use the medians as benchmarks.
(b) As the median values are generally higher for associate than for full professors, the large majority of assistant professors can apply for both associate and full professorship.
(c) This analysis estimates that about 20% (30%) of current assistant (associate) professors are eligible for higher academic positions. These results do not consider candidates who are not currently employed by an Italian University in one of the three SSDs or the cross and multiple applications among the three different SSDs.[23]

The next empirical exercise is aimed at investigating whether the set of eligible candidates is affected by the choice of the bibliometric archive. Table 6 presents the number and the percentages in their respective SSD as well as the academic roles of candidates who are eligible according to both Scopus and WoS (Scopus=WoS), only by Scopus, only by WoS, and candidates who are eligible in at least one of the two archives (Scopus and/or WoS).

Table 6: Differences at the subjective level between eligible candidates using Scopus and WoS database

	Scopus=WoS		**Only Scopus**		**Only WoS**		**At least 1**	
From:	**Assoc.**	**Full**	**Assoc.**	**Full**	**Assoc.**	**Full**	**Assoc.**	**Full**
p/01								
Assistant Prof.	47	39	14	37	18	11	79	87
Associate Prof.	-	64	-	24	-	6	-	94
p/02								
Assistant Prof.	22	14	14	19	1	3	37	36
Associate Prof.	-	27	-	16		1	-	44
p/03								
Assistant Prof.	8	9	6	6	1	2	15	17
Associate Prof.	-	7	-	5	-	2	-	14
Total								
Assistant Prof.	77	62	34	62	20	16	131	140
Associate Prof.	-	98	-	45	-	9	-	152

This table points out how there are relevant differences at the subjective level to select candidates using Scopus instead of WoS data. For instance, eligible candidates according to both archives are:
- 77 on 111 by Scopus (or 97 by WoS) of eligible assistant professors for associate professorship;
- 62 on 124 by Scopus (or 78 by WoS) of eligible assistant professors for full professorship;
- 98 on 104 by Scopus (or 107 by WoS) of eligible associates for full professorship.

[23] According to the different medians among SSDs, we would expect that some assistant and associate professors of p/01 that do not pass the higher benchmarks of their SSD, may apply for associate and/or full professorship to p/02 or p/03.

3.2 CUN eligibility criteria for associate and full professorships

The CUN proposes different criteria according to the specific area of research and academic qualification. By focusing on area 13 (Economics and Statistics), in order to be admitted to the national peer-review evaluation, the economists (p/01, p/02, and p/03) must hold each of the following necessary benchmarks.

To be eligible for peer-review evaluation for associate professorship:

- 1° criterion: The admitted candidate should be (co-)author, at least, of an average of 0.8 publications per year for every year after PhD graduation. The latter average should be calculated accounting for the sole scientific products (pre-PhD included) published under a transparent and documented peer-review.
- 2° criterion: The eligible candidate should have at least six publications as defined by the previous criterion.
- 3° criterion: It is necessary that at least two publications are published in international journals, or publishers of well known scientific value, over the past five years.

To be admitted for peer-review evaluation to become full professor:

- 1° criterion: The admitted candidate should be (co-)author, at least, of an average of one publication per year for every year after PhD graduation. The latter average should be calculated accounting for the sole scientific products (pre-PhD included) published under a transparent and documented peer-review.
- 2° criterion: The eligible candidate should have at least ten publications as defined by the previous criterion.
- 3° criterion: It is necessary that at least four publications are published in international Journals, or publishers of well known scientific value, over the past six years.

In order to be effectual, the CUN criteria need to be integrated with at least three discretionary hypotheses concerning: (1) the year of PhD graduation; (2) the number of peer-review scientific products; and (3) the list of international journals, or publishers of well-known great scientific value.

On the first point, we refer to the ANVUR proposal for the identification of the academic age of each candidate. In particular, it is assumed that the year of the first publication indexed by EconLit can be considered a reliable proxy of the year of PhD graduation. This assumption has been roughly tested by comparing the CNVSU (2010) data on the age of full, associate, and assistant professors for economic area ("*Area 13*"). It has been estimated by EconLit that associate and full professors have published the first peer-review product about 15 and 23 years ago. By considering that the PhD graduation occurs at about 30 years of age, we calculate that the age of associate and full professors should be 45 and 53 years

old, respectively. In fact, these values are comparable to the average age of 48 and 54 years, respectively, drawn from the CNVSU (2010) database.[24]

On the second discretional choice, the CUN (2011) does not cite a specific bibliometric archive for the extraction of the scientific performances of each scholar. Instead, we opt for the widest international bibliometric database for the economic literature: EconLit. In contrast with the two databases recommended by ANVUR (Scopus and WoS), EconLit has the advantage of indexing the main Italian journals (46)[25], and therefore, it does not prejudice the bibliometric performances of scholars who publish mainly in Italian. Further, the use of this database for the evaluation of Italian economists has also been supported by Gagliarducci et al. (2005), Caianelli et al. (2006), Perotti (2008) and Marcuzzo and Zacchia (2007, 2010).

Concerning the third discretional hypothesis, we presume that journals indexed by Scopus can be recognized as international journals of well-known great scientific value. It could be argued that this assumption is not binding enough; however, as shown in Table 7, a more binding criterion (e.g. a list of top-journals) would make the picking of candidates excessively selective and, as a result, ineffectual for the next peer-review evaluation.[26]

[24] We also estimate the age of assistant professors. In this case, we find that the first peer review product has been published seven years ago, therefore, we expect an age of 30+7= 37 years. However, for assistant professors the average age reported by CNVSU (2010) is higher (47 years old). This inconsistency is partially solved by considering the median value that is posited in the range of 41-45 years old. Further, by using the medians instead of the means for associate and full professors as well, the previous estimates of the time elapsed by the PhD graduation does not significantly change.

[25] This estimate is based on Marcuzzo and Zacchia (2007, p. 282). They estimated that the 46 Italian journals constitute 3.95% of the total number of journals indexed by EconLit (1215) and 32% of Italian Economic journals indexed by ESSPER (144).

[26] Alternatively, top-journal requirements may be used in order to define the third criterion compatibly with the introduction of less stringent quantitative benchmarks for the former two CUN criteria.

Table 7: Eligible candidates according to CUN criteria

To:	**Associate Professor**				**Full Professor**			
From:	**1°crit.**	**2°crit.**	**3°crit.**	**Eligible**	**1°crit.**	**2°crit.**	**3°crit.**	**Eligible**
p/01								
Assistant Prof. (310)	136 (44%)	85 (27%)	168 (54%)	**61 (20%)**	123 (40%)	25 (8%)	106 (34%)	**21 (7%)**
Associate Prof. (211)	-	-	-	**-**	47 (22%)	73 (35%)	86 (41%)	**32 (15%)**
p/02								
Assistant Prof. (128)	54 (42%)	37 (29%)	59 (46%)	**28 (22%)**	50 (39%)	14 (11%)	41 (32%)	**13 (10%)**
Associate Prof. (95)	-	-	-	**-**	19 (20%)	30 (32%)	37 (39%)	**16 (17%)**
p/03								
Assistant Prof. (69)	23 (33%)	15 (22%)	29 (42%)	**13 (19%)**	20 (29%)	3 (4%)	15 (22%)	**3 (4%)**
Associate Prof. (37)	-	-	-	**-**	6 (16%)	7 (19%)	8 (22%)	**3 (8%)**
Total								
Assistant Prof. (507)	213 (42%)	137 (27%)	256 (50%)	**102 (20%)**	193 (38%)	42 (8%)	162 (32%)	**37 (7%)**
Associate Prof. (343)	-	-	-	**-**	72 (21%)	110 (32%)	131 (38%)	**51 (15%)**

The following conclusion can be made from Table 7:

a) The most selective benchmarks for the CUN eligibility system are: (i) the second criterion for assistants who wish to be directly enrolled as full professors (i.e. at least ten peer-review publications), and (ii) the first criterion for associate professors (at least one publication per year for every year after PhD graduation).

b) As the benchmarks are the same for the three SSDs, while scientific productivity is different among these sub-disciplines, the CUN selection criteria might be excessively binding for full professorship in p/03.

(c) This analysis estimates that about 20% (15%) of current assistant (associate) professors are eligible for upper academic positions.[27]

In conclusion, by comparing ANVUR and CUN criteria, we deduce that both selection systems are effective in reducing the set of eligible candidates. In addition, the shares of candidates admitted to national peer-review evaluation for associate professorship seem to be very similar. On the contrary, the picking for full professorship is definitely more selective whenever CUN criteria are applied.

These results may certainly change according to the different approaches adopted to tackle (i) the 4th criterion (ANVUR) and/or (ii) the estimation and normalization of academic age (ANVUR), (iii) the year of PhD graduation (CUN), (iv) the number of peer-review scientific products (CUN), and (v) the list of international journals of "great scientific value" (CUN).

[27] Similar to the previous results based on ANVUR criteria, these outcomes neither consider candidates who are not currently employed by an Italian University in one of the three SSDs nor the cross and multiple applications among the three different SSDs.

The next analysis is aimed at verifying if, at the subjective level, the eligible candidates are the same according to the CUN and ANVUR criteria. Tables 8 and 9 present the number and the percentages of eligible scholars according to both CUN and ANVUR based on Scopus and WoS archives, respectively.

Table 8: Differences in predicted career progressions, based on CUN criteria and on the Scopus database

	CUN=Scopus		**Only CUN**		**Only Scopus**		**At least 1**	
	Assoc.	**Full**	**Assoc.**	**Full**	**Assoc.**	**Full**	**Assoc.**	**Full**
p/01								
Assistant Prof.	33 (37%)	19 (24%)	28 (31%)	2 (3%)	28 (31%)	57 (73%)	89	78
Associate Prof.	-	31 (35%)	-	1 (1%)	-	57 (64%)	-	89
p/02								
Assistant Prof.	25 (64%)	13 (39%)	3 (8%)	0 (0%)	11 (28%)	20 (61%)	39	33
Associate Prof.	-	15 (34%)	-	1 (2%)	-	18 (41%)	-	44
p/03								
Assistant Prof.	9 (50%)	3 (20%)	4 (22%)	0 (0%)	5 (28%)	12 (80%)	18	15
Associate Prof.	-	3 (25%)	-	0 (0%)	-	9 (75%)	-	12
Total								
Assistant Prof.	67 (46%)	35 (28%)	35 (24%)	2 (2%)	44 (30%)	89 (71%)	146	126
Associate Prof.	-	49 (34%)	-	2 (1%)	-	84 (58%)		145

Table 9: Differences in predicted career progressions, based on CUN criteria and on the WoS database

	CUN=WoS		**Only CUN**		**Only WoS**		**At least 1**	
	Assoc.	**Full**	**Assoc.**	**Full**	**Assoc.**	**Full**	**Assoc.**	**Full**
p/01								
Assistant Prof.	30 (31%)	15 (27%)	31 (32%)	6 (11%)	35 (36%)	35 (63%)	96	56
Associate Prof.		27 (36%)		5 (7%)		43 (57%)		75
p/02								
Assistant Prof.	17 (50%)	7 (30%)	11 (32%)	6 (26%)	6 (18%)	10 (43%)	34	23
Associate Prof.		13 (42%)		3 (10%)		15 (48%)		31
p/03								
Assistant Prof.	6 (38%)	3 (27%)	7 (44%)	0 (0%)	3 (19%)	8 (73%)	16	11
Associate Prof.		3 (33%)		0 (0%)		6 (67%)		9
Total								
Assistant Prof.	53 (36%)	25 (28%)	49 (34%)	12 (13%)	44 (30%)	53 (59%)	146	90
Associate Prof.		43 (37%)		8 (7%)		64 (56%)		115

According to tables 8 and 9, it can be mentioned that there are relevant differences at the subjective level to choose CUN as to ANVUR benchmarks for the eligibility of candidates. For instance, out of 102 (37) assistant professors eligible for associate (full) professorship according to CUN criteria about 60 (30) are the same as the ANVUR criteria; also, out of 51 associates eligible for full professorship by CUN benchmarks about 45 are the same according to ANVUR eligibility criteria.

It is found that with exclusion of the associate professorship, there are minor differences in eligible candidates selected by CUN criteria and by ANVUR criteria. On the contrary, relevant differences occur when comparing candidates selected by ANVUR with respect to CUN. This is due to the greater number

of eligible candidates who hold ANVUR's benchmarks (378 by Scopus; 282 by WoS) with respect to those holding CUN's criteria (190).

Finally, Table 10 demonstrates at the subjective level whether the eligible candidates differ according to different combinations of CUN and ANVUR criteria.

Table 10: Eligible candidates according to CUN and ANVUR (Scopus and Wos)

	CUN = WoS= Scopus		**Hold 2 criteria**		**At least 1 criterion**		**None of the criteria**	
	Assoc.	**Full**	**Assoc.**	**Full**	**Assoc.**	**Full**	**Assoc.**	**Full**
p/01								
Assistant P. (310)	27 (9%)	14 (5%)	29 (9%)	31(10%)	48 (15%)	43 (14%)	206 (66%)	222 (72%)
Associate P. (211)		26 (12%)		44(21%)		24 (11%)		117 (55%)
p/02								
Assistant P. (128)	17 (13%)	7 (5%)	13 (10%)	13(10%)	10 (8%)	16 (13%)	88 (69%)	92 (72%)
Associate P. (95)		12 (13%)		19(20%)		13 (14%)		51 (54%)
p/03								
Assistant P. (69)	6 (9%)	3 (4%)	5 (7%)	6(9%)	8 (12%)	8 (12%)	50 (72%)	52 (75%)
Associate P. (37)		3 (8%)		4(11%)		7 (19%)		23 (62%)
Total								
Assistant P. (507)	50 (10%)	24 (5%)	47 (9%)	50(10%)	66 (13%)	67 (13%)	344 (68%)	366 (72%)
Associate P. (343)		41 (12%)		67(20%)		44 (13%)		191 (56%)

It is found from Table 10 that:

(a) 68% (56%) of assistant (associate) professors are not eligible with any criteria for associate (full) professorship;

(b) 10% (12%) of assistant (associate) professors are eligible for associate (full) professorship irrespective of the criteria applied.

4. ANVUR and CUN eligibility criteria for commissioners

To be eligible in one of the four positions[28] of commissioner for the national panel for scientific qualification, a person must be: a) full professor for an Italian University; b) hold the (pre-) requisite of scientific productivity required to apply for full professorship in that SSD; and c) submit formal application to MIUR to be included in the list from where the examiners will be drawn.

In the following two sub-sections, details on the two systems for selecting full professors, interested in applying for the national panel of experts, are provided.

4.1 ANVUR eligibility criteria for commissioners

ANVUR criteria of eligibility for examiner expect the aspirant commissioner to hold the requirement needed to apply for full professorship. Nevertheless, there is just a change about the first criterion: the number of publications must not be normalized for academic age.

[28] The fifth component of the panel will be drawn from a list of foreign scholars and experts proposed by ANVUR.

As a consequence, the applicant must hold his/her number of publications indexed by WoS or Scopus[29] in the last ten years, and the number of citations and h-index above the median of full professors in their SSD (Table 11). Similar to applicants for full professorship, the aspirant examiner must also hold the 4th criterion of "reasonable" continuity in his/her scientific production.

Table 11: Medians according to ANVUR criteria estimated using Scopus and WoS (2011) - examiners

	Full Professors (Scopus)			Full Professors (WoS)		
	1°crit. (# publ.)	**2°crit. (# cit.)**	**3°crit. (h-)**	**1°crit. (# publ.)**	**2°crit. (# cit.)**	**3°crit. (h-)**
p/01	5.0	5.0	2.0	2.0	3.3	2.0
p/02	4.0	4.0	1.0	1.0	1.7	1.0
p/03	1.0	1.0	1.0	0.0	0.0	1.0

Table 12 reports the number (share) of full professors eligible for the list of qualified examiners for each SSD.

Table 12: Eligible examiners according to ANVUR criteria (Scopus and WoS)

To:	**Full Professor (Scopus)**					**Full Professor (WoS)**				
From:	**1°crit.**	**2°crit.**	**3°crit.**	**4°crit.**	**Eligible**	**1°crit.**	**2°crit.**	**3°crit.**	**4°crit.**	**Eligible**
p/01										
Full Professor (286)	133 (47%)	130 (45%)	127 (44%)	151 (53%)	**128 (45%)**	142 (50%)	136 (48%)	110 (38%)	120 (42%)	**120 (42%)**
p/02										
Full Professor (116)	57 (49%)	47 (41%)	54 (47%)	52 (45%)	**51 (44%)**	54 (47%)	57 (49%)	45 (39%)	30 (26%)	**30 (26%)**
p/03										
Full Professor (75)	35 (47%)	37 (49%)	30 (40%)	23 (31%)	**23 (31%)**	33 (44%)	33 (44%)	20 (27%)	17 (23%)	**17 (23%)**
Total										
Full Professor (477)	225 (47%)	214 (45%)	211 (44%)	226 (47%)	**202 (42%)**	229 (48%)	226 (47%)	175 (37%)	167 (35%)	**167 (35%)**

Based on the results of Table 12, it can be mentioned that scientific production benchmarks estimated by Scopus database are slacker than those of WoS.

By considering the three SSDs as a whole, about 60% of full professors are not eligible for the role of examiner to the national panel for scientific qualification. Likewise, for the eligibility of candidates, in this case as well, the ANVUR system may fine-tune the sample of eligible examiners by setting a binding criterion of "reasonable" scientific continuity.

4.2 CUN eligibility criteria for commissioners

Similar to the ANVUR proposal, CUN criteria of eligibility as examiner require the aspirant evaluator to hold the requisites essential to apply for full professorship. The aspirant examiner must be (co-)author of at least one peer-reviewed publication per year for every year after PhD graduation, must demonstrate at

[29] Or any other database that has characteristics of wide coverage and methodological rigor.

least ten peer-review publications and, finally, must be (co-)author of four publications, at least, issued by international journals or publishers of well known great scientific value over the past six years.

Table 13: Eligible commissioners according to CUN criteria

To:	Commissioner			
From:	**1°crit.**	**2°crit.**	**3°crit.**	**Eligible**
p/01				
Full Professor (286)	77 (27%)	192 (67%)	120 (42%)	**64 (22%)**
p/02				
Full Professor (116)	22 (19%)	65 (56%)	37 (32%)	**17 (15%)**
p/03				
Full Professor (75)	12 (16%)	29 (39%)	17 (23%)	**8 (11%)**
Total				
Full Professor (477)	111 (23%)	286 (60%)	174 (36%)	**89 (19%)**

Table 13 reveals that the benchmarks proposed by CUN are more binding than ANVUR benchmarks. Indeed, about 81% of full professors are not eligible for the role of examiner to the national panel for scientific qualification. Most of the selection is a result of the joint effect of the first criterion (at least one publication per year for every year after PhD graduation) and the advanced academic age of Italian full professors.

Table 14 is aimed at verifying whether the eligible examiners are the same according to the CUN and ANVUR criteria. The number and the percentages of those eligible according to different combinations of eligibility criteria are shown below.

Table 14: Eligible examiners according to CUN and ANVUR (Scopus and Wos)

	CUN = WoS = Scopus	**CUN = Scopus**	**CUN = WOS**	**Scopus = WoS**	**At least 1 criterion**	**None of the criteria**
p/01						
Full Prof. (286)	25 (9%)	30 (10%)	25 (9%)	114 (40%)	57 (20%)	110 (38%)
p/02						
Full Prof. (116)	11 (9%)	14 (12%)	12 (10%)	28 (24%)	33 (28%)	51 (44%)
p/03						
Full Prof. (75)	2 (3%)	3 (4%)	3 (4%)	14 (19%)	20 (27%)	39 (52%)
Total						
Full Prof. (477)	38 (8%)	47 (10%)	40 (8%)	156 (33%)	110 (23%)	200 (42%)

It is found from Table 14 that:

(a) About 40% of full professors are not eligible with any criteria for the role of examiner;

(b) Only 8% of current full professors are eligible as examiner irrespective of the criteria applied.

5. How is candidate selection expected to change with ANVUR criteria? A simulation for academic positions advertised in 2005

We evaluate if, and how, the profiles of the "next" generation of academic economists may change according to the reformed recruitment procedures. For the nature of simulation, these results do not take

into account the expected behavioral responses of upcoming candidates in terms of their editorial strategies (e.g. incentives to publish on Scopus or WoS indexed journals, to use self-citations, co-authorship).

To develop our arguments, we envisage asking ourselves the following questions:

1) How robust are the results of academic competitions in 2005 with respect to the introduction of ANVUR eligibility criteria for candidates?
2) How robust is the composition of local commissions in 2005 with respect to the introduction of ANVUR eligibility criteria for examiners?
3) Does the previous recruitment system (L.210/1998) forecast candidates' bibliometric performances better than the ANVUR benchmark system (L.240/2010)?

To answer these questions empirically, a new *ad hoc* database has been constructed that includes all the public local competitions for positions as full and associate professorships in the area of Economics (p/01, p/02, p/03) advertised from the 4th session of competitions held in 2004 to the 4th session of 2005 as resulting from the MIUR website.

The year 2005 has been chosen as it is the most recent year with a satisfactory number of academic recruitments. In fact, the number of *idonei* nominated by the entire academic recruitment system is as follows (CNVSU, 2010): 2002 - full (1168), associate (1929); 2003 - full (639), associate (1008); 2004 - full (1031), associate (1545); 2005 - full (1174), associate (1857); 2006 - full (191), associate (234); 2007 - full (3), associate (3); 2008 - full (377), associate (572); 2009, 2010 - full (0), associate (0).

The strong reduction of *idonei* after 2005 is due to both the decrease of advertised positions (*bandi*) and the limitation to the maximum number of winners that each local competition could declare. In particular, for each *concorso* the commission could declare up to: 3 *idonei* since 1999 to 2001; 2 *idonei* between 2002 and May 2005; 1 *idoneo* between May 2005 and May 2006; 1 *idoneo* since May 2006 up to the second round (*tornata*) of 2008 (with the exclusion of 2 *idonei* for the first *tornata* of 2008). Since the second round of 2008 the academic recruitment has been suspended in order to await the definition and approval of the new recruitment system.

The *ad-hoc* sample includes 45 competitions with two *idonei* and 4 competitions with one *idoneo* (University of Messina: p02, Kore University of Enna; p01, University of Perugia: 2 of p-03).

The list of "effective" participants has been extrapolated from the final official report (*verbale finale*), that is, the sole candidates evaluated by the examiners have been considered. As a result of very stringent data protection policies of some Universities, 11 out of the 49 competitions in the reference period have been excluded from the analysis of applicants' performances;[30] however, they have been considered for the evaluation of commissioners.

[30] For 27 out of 49 competitions, the official reports have been downloaded from the recruitment office website. For 11 competitions, the list of effective participants has been formally requested and obtained from the recruitment office under precise privacy statements. For the remaining competitions, our request has been rejected (2

This analysis has been simulated by applying only ANVUR criteria on the Scopus database. On the one hand, as both CUN thresholds are not sample-based and the candidate's bibliometric scores have usually increased in the recent years,[31] CUN benchmarks may be excessively selective to be applied to these competitions. On the other hand, with respect to the ANVUR picking system, the Scopus dataset has been preferred, owing to the empirical evidence that: (i) it has a wider coverage of publications and citations for economic literature than WoS, and (ii) there are no significant differences between the outputs achieved by the archives.

That being stated, bibliometric and citational performances are extracted for each of the 854 academic professors (350 associate and 504 full professors) employed by Italian Universities dated 31/12/2004. Table 15 reports the median values required for estimating the first three ANVUR criteria for the eligibility of candidates.

Table 15: Medians according to ANVUR criteria estimated using Scopus database (2004) - candidates

	Associate Professors			Full Professors		
	1°crit. (# publ.)	**2°crit. (# cit.)**	**3°crit. (h-)**	**1°crit. (# publ.)**	**2°crit. (# cit.)**	**3°crit. (h-)**
p/01	2.0	0.0	0.0	2.0	2.0	1.0
p/02	0.0	0.0	0.0	0.0	0.0	0.0
p/03	2.2	0.0	0.0	0.0	0.0	0.0

Concerning the second necessary requirement (4th criterion), it is assumed that a "reasonable" continuity in the scientific production for a candidate applying for associate (full) professorship in the 2005 occurs whenever at least one (two) international articles indexed by Scopus over the period 2000-2004 (1995-2004) has been produced. These values are half of the thresholds applied in 2011. Such a slackening of the benchmark is mainly motivated by the empirical evidences on the significant escalation of bibliometric performances in the last decade.[32] In particular, by applying for 2005 the same benchmarks employed for 2011, 44% (186 on 424) of candidates would be eligible to fill the 72 vacancies; moreover, 54% of winners (39) would not be eligible. As a result, it may be observed that such a strong picking may excessively unbalance the system towards "quantitative" criteria, therefore reducing the fundamental role of peer-review evaluation to select among candidates.[33]

competitions) or disregarded (9 competitions). The 11 missing competitions consist of 8 associate p/01 professorships, 2 full p/01 professorships, and 1 associate p/02 professorship.

[31] Based on the EconLit database we estimate that 44.0% (25.5%) of associate (full) professors' publications have been published later than 2005.

[32] By looking exclusively at the 424 examined candidates, we estimate that the items published since 2005 are twice (194.6%) the publications issued over the period 2002-2004.

[33] Details on this analysis are available from the authors upon request.

5.1 How robust are the results of academic competitions in 2005 with respect to the introduction of ANVUR eligibility criteria for candidates?

Following the evaluations made by local panels, the 424 applicants of academic competitions for professorship in 2005 are grouped into two groups (winners and losers). For each group we detect how many *idonei* and losers may have been eligible for associate and full professorships according to ANVUR criteria (Table 16).

Table 16: Overall results of 38 Local Competitions for professorship in 2005

	# competitions	# applicants	# positions	Winners Eligible (240/2010)	Winners No Eligible (240/2010)	Losers Eligible (240/2010)	Losers No Eligible (240/2010)
Associate Professors							
p/01	10	80	22	9 (11%)	13 (16%)	17 (21%)	41 (51%)
p/02	6	53	12	3 (6%)	9 (17%)	9 (17%)	32 (60%)
p/03	1	2	1	0 (0%)	1 (50%)	0 (0%)	1 (50%)
Total	**17**	**135**	**33**	**12 (9%)**	**23 (17%)**	**26 (19%)**	**74 (55%)**
Full Professors							
p/01	15	232	27	16 (7%)	11 (5%)	106 (46%)	99 (43%)
p/02	3	24	5	4 (17%)	1 (4%)	9 (38%)	10 (42%)
p/03	3	33	5	1 (3%)	4 (12%)	12 (36%)	16 (48%)
Total	**21**	**289**	**37**	**21 (7%)**	**16 (6%)**	**127 (44%)**	**125 (43%)**
Total							
p/01	*25*	*312*	*49*	*25 (8%)*	*24 (8%)*	*123(39%)*	*140 (45%)*
p/02	*9*	*77*	*17*	*7 (9%)*	*10 (13%)*	*18 (23%)*	*42 (55%)*
p/03	*4*	*35*	*6*	*1 (3%)*	*5 (14%)*	*12 (34%)*	*17 (49%)*
Total	***38***	***424***	***72***	***33 (8%)***	***39 (9%)***	***153 (36%)***	***199 (47%)***

Notes in parenthesis are percentages with respect to the number of applicants.

It can be deduced from Table 16 that:

(a) 54% of winners (39 on 72) were not eligible for the academic role according to ANVUR bibliometric criteria.

(b) By applying ANVUR benchmarks to select candidates, the recruitment procedures would have filled the 72 vacancies by peer-reviewing only 186 candidates, instead of 424. As a result, 39 (not eligible) winners would have been replaced by some of the 153 eligible losers.

Nevertheless, for the 11 competitions with missed final reports (22 professorships) we cannot report the comparative analysis of their final decisions. For these *concorsi*, the examination is only based on bibliometric data about the winners (because they are available from MIUR database). According to these results, for 16 (4) *idoneità* as associate (full) professor of p/01, 56% (75%) of winners were not eligible for higher academic roles. On the contrary, for the competition for associate professorship of p/02, both winners were eligible according to ANVUR criteria.

The second simulative exercise involves evaluating whether a non-eligible winner may be replaced by at least one eligible loser. According to the rules of the previous system, this analysis is carried out by

separately comparing winners and losers for each local competition. The simulation reveals that for 45% of local competitions (21 on 38), the commission would have declared at least one different winner.

Thus, we conclude that the introduction of ANVUR criteria will remarkably change the academic recruitment system; in particular, it is expected to effectively increase scientific productivity (ANVUR scores) of the winners.

5.2 How robust is the composition of local commissions in 2005 with respect to the introduction of ANVUR eligibility criteria for examiners?

This section is aimed at assessing how the composition of local commissions would have changed by applying ANVUR minimum standards for the eligibility of examiners.

This sample includes 49 local commissions each with 5 examiners (245 commissioners).

First, it is clear to these authors that it is not possible to compare exactly the old and new rules of panel compositions. In particular, according to L.240/2010 it will be sufficient to select 15 examiners (four Italians full professors and a foreign well-known scholar drawn from a list proposed by ANVUR for each SSD) instead of 245 professors required for evaluating the candidates for 2005.

The second relevant difference between the incoming structure of the panel of experts and the previous local commission is that for associate professorship it was also required to include two associate professors among the five members of the examining board. It implies that this simulation excludes all the 54 associate professors although some of these hold ANVUR benchmarks for full professorship.

Table 17 reports the median values - estimated using Scopus – according to ANVUR criteria for examiners dated 31/12/2004. For completeness, we also include medians for associate professors although they are not relevant to selecting eligible commissioners.

Table 17: Medians according to ANVUR criteria estimated using Scopus data (2004) - examiners

	Associate Professors			Full Professors		
	1°crit. (# publ. no Norm.)	**2°crit. (# cit.)**	**3°crit. (h-)**	**1°crit. (# publ. no Norm.)**	**2°crit. (# cit.)**	**3°crit. (h-)**
p/01	1.0	0.0	0.0	1.0	2.0	1.0
p/02	0.0	0.0	0.0	0.0	0.0	0.0
p/03	1.0	0.0	0.0	0.0	0.0	0.0

Table 18 indicates how many examiners would have again been members of the local commissions in the presence of ANVUR criteria for eligibility in 2005.

Table 18: ANVUR criteria for examiners' eligibility in 2005

# commissioners		1°crit. (# publ.)	%	2°crit. (# citat.)	%	3°crit. (h-)	%	4°crit. (cont.)	%	Eligible	%
p/01 (36 competitions)											
Full Prof.	120	67	56%	52	43%	38	32%	67	56%	**51**	**43%**
Ass. Prof.	60	-		-		-		-		**-**	
Total	**180**	**67**	**37%**	**52**	**29%**	**38**	**21%**	**67**	**37%**	**51**	**28%**
p/02 (11 competitions)											
Full Prof.	37	14	38%	10	27%	10	27%	9	24%	**9**	**24%**
Ass. Prof.	18	-		-		-		-		**-**	
Total	**55**	**14**	**25%**	**10**	**18%**	**10**	**18%**	**9**	**16%**	**9**	**16%**
p/03 (2 competitions)											
Full Prof.	7	6	86%	5	71%	5	71%	5	71%	**5**	**71%**
Ass. Prof.	3	-		-		-		-		**-**	
Total	**10**	**5**	**60%**	**5**	**50%**	**5**	**50%**	**5**	**50%**	**5**	**50%**
Total (49 competitions)											
Full Prof.	164	87	53%	67	41%	53	32%	81	49%	**65**	**40%**
Ass. Prof.	81	-		-		-		-		**-**	
Total	**245**	**87**	**36%**	**67**	**27%**	**53**	**22%**	**81**	**33%**	**65**	**27%**

According to these statistics, the large majority of examiners involved in recruitment procedures were not sufficiently qualified according to the benchmarks proposed by ANVUR. In particular, only 65 professors who took part in these procedures could have an active role, also in accordance to the rules of the new academic recruitment system.

An additional simulative exercise is aimed at estimating how many local commissions could have changed at least three examiners. This aspect is relevant as if unanimous consent is not achieved then the "*idonei*" are picked by ranking the candidates in order of examiner's preference(s). On this issue, we find that 84% (41 on 49) of competitions have at least three non-eligible examiners.[34]

In the existing literature, a positive correlation is usually conjectured between the bibliometric performances of examiners and their propensity to declare *idonei* the candidates with the best bibliometric performances (Perotti, 2002b, 2003, 2004, 2006). Given the ANVUR criteria for the identification of the "most qualified candidates" (best bibliometric performances), we attempt to empirically test this hypothesis for the 38 academic competitions of our sample.

Our analysis provides evidence of significant correlation between the (not) eligibility of winners and the (not) eligibility of the majority of panel examiners. The statistical test is aimed at calculating the Pearson correlation between two dichotomous variables: the first one assumes value 1 if at least 3 members of examiners are not eligible according to ANVUR qualification requirements and 0 otherwise; the second one assumes value 1 if the commission opted for, at least, a non-eligible winner although there was at least an eligible candidate and 0 if all winners were eligible or there were no eligible candidates among applicants.

[34] These 8 eligible majority in the panel are in 7 *concorsi* for full professorship of p/03 and 1 competition for full professorship of p/01.

The estimated correlation coefficient indicates a significant (positive) relationship between the low qualification of the majority of examiners and the selection of candidates not consistent with the observable criteria of scientific productivity (r^2=0.75). Furthermore, a t-test to check for the statistical significance of correlation is computed. The t-statistic value is 6.77 (p-value 0.000); therefore, the null hypothesis (no correlation) can be rejected.

Thus, we conclude that the ANVUR eligibility criteria for examiners will definitely bring about a change in who decides on academic career. In line with the previous statistical test, this modification might be beneficial in selecting the candidates with the best bibliometric performances.

5.3 Does the previous recruitment system (L.210/1998) forecast candidates' bibliometric performances better than the ANVUR benchmarks system (L.240/2010)?

To our knowledge, there is no literature that has attempted to evaluate the capability of local commissions to forecast the best (bibliometrically) productive scholars among the applicants of each *concorso*. In contrast with the previous literature (e.g. Perotti, 2008), in this section we analyze the after-*concorso* bibliometric performances of the candidates and not (only) the data available to the examiners at the time of the *concorso*. This analysis is made possible by an *ex-post* perspective, where each applicant (winners and losers) of each *concorso* is scrutinized and further compared with the colleagues of competitions by bibliometric data after six years of his/her evaluation. In this way, we aim to assess whether the two-step recruitment system performs better than the previous local peer-review evaluation of candidates.

According to the authors, this kind of analysis is not able to take into account the moral hazard of the winners for full professorship, who after achieving *idoneità* have less incentive to continue their scientific production with respect to the losers. At the same time, a disincentive to the "publish or perish" may also occur for qualified "losers" as they may be discouraged from investing their time to publish by suffering the lack of meritocracy of the local recruitment system.

To conduct this analysis, the bibliometric scores of the 424 applicants of *concorsi* advertised in 2005 are computed twice: at the time of *concorso* (31/12/2004) and on December 2011. The performances of losers and winners are accounted for by using the three ANVUR indicators of scientific productivity (number of publications - not normalized for academic age[35]-, number of citations, and h-index).

In order to focus on the differences between the previous (one-step: peer-review) and the incoming (two-step: bibliometric picking and peer-review limited to eligible candidates) recruitment system, we examine the 21 local *concorsi* where the result of the one-step system conflicts with that of the two-step system, that is, the local commission preferred to declare *idoneo* a not eligible candidate although, among the applicants of that *concorso*, there was at least one eligible candidate.

[35] As of 31/12/2004, all items indexed by Scopus, published between 1995 and 2004, are accounted for. For the post-competition analysis of scientific productivity only the items published since 2005 are considered.

For each of these competitions, the three indicators of bibliometric performances (#pub., #cit. and h-) of the eligible losers (EL) are compared with those of non-eligible winners (NEW).

It is assumed that EL is more scientifically productive than NEW if at least two of the three scientific productivity scores held by an EL are greater than the NEW score. In this case, we assert that the one-step system underperforms the two-step recruitment system.

According to these definitions, by analysing the candidates' productivity between 2005 and 2011, the probability that an EL outperforms a NEW is 0.78 (93 times on 119). In 26 cases a NEW shows scientific records better than an EL. Nevertheless, with the exception of 1 *concorso* (for associate professorship in p/01), in 95% of public competitions (20 on 21) there is at least an EL that outperforms the NEW.[36]

Thus, we conclude that the ANVUR criteria of eligibility for candidates may be beneficial in selecting the candidates with the best bibliometric performances and, mostly, in reducing the possibilities for manipulation and pressure in the academic recruitment system.

6. Some hints for ANVUR criteria

It is quite relevant to point out that the median approach proposed by ANVUR is expected to become increasingly more selective in the following years owing to at least three factors: retirements of personnel with lower bibliometric scores; the dynamic effect of benchmarks based on the medians; and the change of editorial strategy of the academic staff.

On average, older academic personnel are usually characterized by lower bibliometric scores than younger colleagues. This is probably owing to the different editorial strategy followed in the past decades (e.g. greater relevance of books and articles published in Italian journals) as well as the different beliefs on the main objectives of academic research (e.g. overspecialization). As a result, the next career promotions are likely to increase the gap among median values at different career levels. The *abilitati* will have bibliometric scores above the median values of the incoming professorship, therefore they will increase the incoming medians progressively. At the same time, the candidates who move up in the academic hierarchy would lower the median value of their previous role. The speed of the adjustment towards the new equilibrium will depend on several factors: (1) the impact of the bibliometric scores for the peer-review evaluation of the national panel; (2) the role assigned to bibliometric scores for the selection at the local level of the scholars to hire among the eligible applicants; and (3) the effective possibilities for the Universities to engage (or promote in the higher position) the *abilitati.*[37]

The incoming system of qualifications for academic competition is likely to change the editorial strategy of scholars. Intuitively, the first best strategy for scholars with insufficient bibliometric performances

[36] Of course, there is no guarantee that the peer-review selection provided by L. 240/2010 will be able to choose exactly that candidate.

[37] For instance, changes in the allocation of FFO ("*Fondo di Finanziamento Ordinario" - Ordinary Financing Fund*) strongly affect the recruitment policy of the Italian University system. At this time, that excludes the Universities with an amount of personnel costs above 90% of FFO to engage new professors.

consists of increasing, firstly, the number of articles published on Scopus or WoS indexed journals (possibly with self-citations), and subsequently, once the eligibility criteria are satisfied, to aim mainly for leading journals.

According to these results, we expect that if national competitions would be advertised yearly as stated by the L. 240/2010, in a relatively few years, the composition of Italian academic staff would change. It would increase the gap among medians making the selection considerably more stringent, especially for discrimination among associate and full professors.

On the basis of our evidence, we point out some hints to potentially improve ANVUR criteria for selecting candidates and commissioners.

About the selection of the database: since the bibliometric rankings of researchers differ at the subjective level depending on the database, the identification of accredited publications/citations should be better motivated on the base of largely agreed (and transparent) criteria. Alternatively, in the absence of an agreement, the eligibility criteria for candidates and commissioners should be intended as met whenever this result occurs for at least one of the reliable and non-manipulable citation databases (e.g., Scopus and WoS).[38]

About the first criterion: we deem the year of PhD graduation as the best choice to normalize scientific production for the academic age. In this sense, we suggest abandoning a normalization anchored to Scopus or WoS data as proposed by ANVUR, or alternatively, using the year of first publication as resulting from EconLit instead of Scopus or WoS. This hint is based on a comparative analysis of the bibliometric profiles of Italian academic economists. For these scholars we estimate the averages of the academic age for the three bibliometric archives. When the EconLit database is considered, the academic age is found to be 7, 15, and 23 years, respectively, for assistant, associate, and full professors. These values are, respectively, 5, 10, 15 years and 4, 9, and 15 years when the first publications from Scopus and WoS are considered. Taking into account that PhD graduation occurs at about thirty years of age, the ages of assistant, associate, and full professors is assumed to be about 37, 45, and 53 years by EconLit; 35, 40, and 55 years by Scopus; and 34, 39, and 45 years by WoS. As reported by CNVSU (2010) these ages are lower than the effective average ages of Italian professors: 47, 48, and 54 years, respectively. In this sense, the comparison among the databases reveals that EconLit provides the best estimates.

About the third criterion: we suggest using the g-index (Egghe, 2006) instead of the h-index (Hirsch, 2005). As observed in the existing literature (e.g. Bornmann, 2007; Jin, 2006; Burgos, 2010), the main

[38] On the basis of our simulations (Tables 6, 8, 9, 10 and 14), this approach increases the shares of candidates admitted to national peer-review evaluation. For instance, applying ANVUR criteria (table 6), we estimate that the share of assistant professors eligible for associate (full) professorship increases of 35% (79%) compared to WoS database and of 18% (13%) compared to Scopus database. Similarly, the share of associate professors eligible for full professorship increases of 42% (6%) compared to WoS (Scopus) database.

cons of the h-index are (I) sensitivity to time passed since the date of publication,[39] (II) invariance with respect to the number of co-authors,[40] (III) excessive sensitivity to variations in the number of publications and/or citations for "not-so-cited" scientific outputs,[41] (IV) unreliability for the comparison of heterogeneous scientific outputs, (V) depreciation of highly cited publications (*seminal papers*),[42] and (VI) depreciation of citation vectors with few but highly cited papers.[43] The first three cons are usually managed through the introduction of *ad hoc* weighting procedures (Katsaros et al., 2006; Batista et al., 2006; Jin et al., 2007; Schreiber, 2008; Wu, 2010; Harzing, 2011). Con (IV), on the other hand, is strictly related with the selection of the research units to be compared at the preliminary stage. Finally, the use of the g-index (Egghe, 2006)[44] may allow overcoming cons (V-VI).[45]

To provide empirical support for the previous hint, Table 19 reports the g- and h-indexes for the 1327 economists of our sample. There is evidence that the g-index provides more differentiation of the citational profiles of the academic roles ($g \geq h$). As a consequence of low medians of h-index scores, it may thus help scientific assessment.

Table 19: Medians of g- and h-index estimated using Scopus and WoS (2011)

	Scopus						**WoS**					
	Assistant		**Associate**		**Full**		**Assistant**		**Associate**		**Full**	
	3° (h-)	3° (g-)	3° (h-)	3° (g-)	3° (h-)	3° (g-)	3° (h-)	3° (g-)	3° (h-)	3° (g-)	3° (h-)	3° (g-)
p/01	1.0	1.0	2.0	2.0	2.0	4.0	0.0	0.0	1.0	2.0	2.0	3.0
p/02	1.0	0.0	1.0	2.0	1.0	2.0	0.0	0.0	1.0	1.0	1.0	1.0
p/03	0.0	0.0	1.0	1.0	1.0	1.0	0.0	0.0	1.0	1.0	1.0	1.0

[39] Evidently, the interval of time for an adequate assessment of the contribution of each publication may strongly differ depending on the social and cultural environment, the discipline and, within the same discipline, on the nature of the contribution.

[40] Co-authorship may assume very different characteristics depending on the nature of the collaboration (Bruno, 2010). As a result, the need for a weighting procedure may be differently supported depending on the presence of opportunistic behaviors.

[41] Let us consider the two citation vectors $x = \{2,1,1\}$ and $y = \{2,2,1\}$ with $h(x) = 1$ and $h(y) = 2$. At the margin, in order to increase $h(x)$, an additional citation for the second publication is sufficient. Differently, in order to induce a marginal increase of $h(y)$, an additional citation for the first and the second publication as well as two additional citations for the third publication are required.

[42] Let us consider the two citation vectors $x = \{1000\}$ and $y = \{2,2\}$ with $h(x) = 1$ and $h(y) = 2$. For the purposes of the h-index, 999 citations in *x* are not relevant at all (excess of citations).

[43] Let us consider two citation vectors $x = \{1000,1000\}$ and $y = \{3,3,3\}$ with $h(x) = 2$ and $h(y) = 3$. As the h-index cannot be larger than the number of publications by construction, *x* is under-estimated with respect to *y* independently of the (evident) larger impact of its scientific output.

[44] Two different versions of the g-index are proposed in Egghe (2006). The former, the score of the g-index is upper bounded by the number of publications. The latter, instead, the score may be reasonably higher than the number of publications in the presence of few but highly cited scientific outputs. As observed Woeginger (2008), the latter "*seems to give the nicest and most natural version of the g-index*".

[45] Alternatively, Abatemarco, Dell'Anno (2012) proposes a generalized class of citation indexes (certainty equivalent citation index) able to overcome cons (V) and (VI).

7. Conclusions

The present research is aimed at assessing how the profile of candidates and examiners may change according to the incoming reformed system designed by Gelmini's reform. There is abundant anecdotal evidence that both the centralized (until 1998) and the local (from 1998-2008) competitions for professorship were prone to manipulation (e.g. Checchi 1999; Perotti 2002, 2008; Durante et al. 2011).

Law 240/2010 is aimed at reducing nepotism and localism of the Italian academic recruitment system. These goals are pursued by centralizing competitions for professorship and, mainly, by selecting candidates and examiners using a set of minimum objective standards of scientific productivity and continuity based on bibliometric data.

It is outside the scope of this study to discuss topics related to the ongoing debate on bibliometric benchmarks: Can scientific production can be measured rather precisely via the bibliometric data? And how? How incentives to research activity may be affected by the bibliometric benchmarks? Even though we did not enter this strand of literature, one may fairly agree that the "price" of preliminary (bibliometric) picking through an imperfect measure of scientific productivity[46] is lesser than the "cost" of a higher risk of opportunistic behaviours in the selection of academic scholars.

Firstly, an empirical analysis is provided for comparing the main bibliometric data sources for Economics. This analysis is propaedeutic to the ongoing debate on the bibliometric scores of candidates. In summary, we find strong empirical evidence that EconLit has the widest coverage of scientific production for Italian academic economists. Looking at the citation archives, Scopus accounts for both more documents and citations than WoS in Economics. However, it is seen that by using sample based statistics (e.g. median values), the use of Scopus and WoS has a slight impact on the shares of candidates admitted to national peer-review evaluation. At the contrary, there are relevant differences at the subjective level to select candidates using Scopus instead of WoS data.

Two main empirical exercises were performed for investigating the consequences of adopting the two alternative picking systems proposed by CUN and ANVUR. The differences between these two sets of criteria reveal a theoretical-based position on the reliability of citation metrics as evaluation criteria to assess research activity. In particular, the CUN proposal consists of absolute thresholds on the quantity of publications normalized by academic age while the ANVUR proposal is characterized by sample based benchmarks (median values) and assigns a relevant role to the citation profile: number of total citations and h-index.

By comparing ANVUR and CUN criteria, we deduce that both picking systems reduce significantly the size of eligible candidates. In particular, the shares of candidates admitted to national peer-review evaluation for associate professorship are very similar. On the contrary, the CUN picking system for full professorship is considerably more selective than the ANVUR system. Of course, these results may

[46] On the limits of bibliometric metrics see Baccini (2010) and Bowles (2007).

change according to different approaches that may be followed to tackle: for the ANVUR system, the 4th criterion or the method to estimate and normalize the academic age; for the CUN system, the ways to determinate the year of PhD graduation, the minimum number of peer-review scientific products, and the list of the international journals of "great scientific value". In this perspective, the descriptions of both systems are (still) lacking. As a consequence, to state which bibliometric picking system has better chances to be manipulation-proof, we would wish for greater operative indications. Once again, the devil may be in the details!

The second simulative exercise includes examining the effects on candidates, commissioners, and quality of decisions if ANVUR eligible criteria had been applied to the local public competitions advertised in 2005. According to our findings, in the absence of respondent behaviour, the recruitment procedure provided by L.240/2010 would strongly affect the selection of the candidate for professorships.

Further, we have also attempted to compare the previous recruitment system (L.210/1998) with the ANVUR benchmark system in terms of their respective capability to forecast candidates' scientific performances. This analysis is done by comparing bibliometric scores between non-eligible winners and eligible losers with respect to the post-*concorso* scientific performances.

According to our findings, a system based uniquely on peer-review evaluation is less proficient for making decisions on career academic progression than a two-step procedure. This is because the two-step system lowers the error of the first kind (false positive: a poor researcher advances to a higher academic position) by reducing the sample of candidates evaluated by peer-review. The national commission has lesser chances to include inadequate researchers in the list of national scientific qualified scholars. On the negative side, this skimming may produce error of the second kind (false negative: a good researcher is not admitted to national peer-review evaluation). That is the "price" of the bibliometric assessment of research. It may be reasonably overcome by using benchmarks criteria that are non excessively binding (e.g. ANVUR by Scopus or CUN with lower thresholds of scientific productivity).

The new recruitment system is essentially expected to reduce the space for manipulation and pressure even if, depending on the size of accountability for decentralised hiring decisions, opportunistic behaviours may still drive the selection of *abilitati* at the University/Department level.

References

Abatemarco A. - Dell'Anno R. (2012). Certainty equivalent citation: generalized classes of citation indexes. Scientometrics (DOI 10.1007/s11192-012-0758-x).

ANVUR (2011a), Criteri e parametri di valutazione dei candidati e dei commissari dell'abilitazione nazionale scientifica. Approvato il 22/06/2011 n. 1/2011. (Disponibile da: http://www.anvur.org/media/277/documento01_11.pdf).

ANVUR (2011b), Sul documento ANVUR relativo ai criteri di abilitazione scientifica nazionale, commenti, osservazioni critiche e proposte di soluzione. Approvato il 25/07/2011 n. 2/2011. (Disponibile da: http://www.anvur.org/sites/anvur-miur/files/Documento%2002_11.pdf).

Baccini A. (2010), Valutare la ricerca scientifica. Uso e abuso degli indicatori bibliometrici, Il Mulino, Bologna.

Bar-Ilan J. (2008), Which h-index? A comparison of WoS, Scopus and Google Scholar, in Scientometrics, vol. 74 (2), pp. 257-271.

Batista P.D. - Campiteli, M.G. and Konouchi, O. (2006), Is it possible to compare researchers with different scientific interests?,Scientometrics, 68(1), pp. 179-189.

Becker S.O. Ichino A. - Peri G. (2004), How Large is the Brain Drain from Italy?. Giornale degli Economisti e Annali di Economia, 63 (1), pp.1-32.

Bonavenutra L. (2011), The Role of Frictions on Italian Academic Recruitment System University of Catania, Department of Economics and Quantitative Methods in its series DEMQ Working Paper Series with number 2011/2.

Bornmann L. (2011), Mimicry in science? Scientometrics, vol. 86, pp. 173-177.

Bruno B. (2010), The Economics of co-authorship, MPRA paper, No. 27730.

Burgos A. (2010), Ranking scientists,Rev. Working Paper Series, 2, Departamento de Fundamentos del Análisis Económico, Universidad de Murcia, Spagna.

Caianelli G. - De Felice A. - Lamonarca M. - Zoboli R. (2006), The publications of Italian Economists in Econlit. Quantitative Assessment and Implications for Research Evaluation, in *Economia Politica*, vol. XXIII (3), pp. 385-423.

Checchi D. (1999), Tenure. An Appraisal of a National Selection Process for Associate Professorship, in Giornale degli economisti e Annali di economia, 58 (2), pp. 137-181.

Bowles S. (2007), La valutazione della ricerca negli Stati Uniti: risultati di un questionario, Rivista Italiana degli Economisti, 12(2), pp. 255-258

CNVSU (2010), Undicesimo Rapporto sullo Stato del Sistema Universitario. Roma

CUN (2011), Criteri, parametri e indicatori per l'abilitazione scientifica nazionale. Sessione n.84 del 24 maggio 2011 (Disponibile da: http://www.cun.it/media/113043/do_2011_05_24_002.pdf).

Durante R. - Labartino G. - Perotti R. (2011), Academic Dynasties: Decentralization and Familism in the Italian Academia, CEPR Discussion Papers 8645.

Egghe L. (2006), An improvement of the h-index: the g-index, ISSI Newsletter vol. 2, pp. 8-9.

Gagliarducci S. – Ichino A. – Peri G., Perotti R. (2005), Lo splendido isolamento dell'universita italiana", in Oltre il Declino, Boeri T- Faini R. – Ichino A. - Pisauro G. – Scarpa C. editors, Il Mulino. Bologna.

Harzing A.W. (2008), Google Scholar - a new data source for citation analysis (disponibile da: http://www.harzing.com/pop_gs.htm).

Harzing A.W. (2010), The Impact of different data sources and citation metrics, (disponibile da: http://www.harzing.com/data_metrics_comparison.htm).

Harzing A.W. (2011), Publish or Perish 3.1,www.harzing.com/pop.htm.

Hicks D. – Wang J. (2009), Toward a Bibliometric Database for the Social Sciences and Humanities, Available at: http://works.bepress.com/diana_hicks/18.

Hirsch J.E. (2005), An index to quantify an individual's scientific research output, in Proceedings of the National Academy of Sciences of the United States of America, vol. 102 (46), pp. 16569-16572.

Jin B.H. - Liang L.M. - Rousseau R. - Egghe, L. (2007), The r- and ar-indices: complementing the h-index, Chinese Science Bulletin, vol. 52, pp. 855-863.

Jin B.H. (2006), H-index: an evaluation indicator proposed by scientist, Science Focus1 (1), pp. 8-9.

Katsaros C. – Manolopoulos Y. - Sidiropoulos A. (2006), Generalized h-index for disclosing latent facts in citation networks, Scientometrics, vol. 72 (2), pp. 253-280.

Labbé C. (2010), Ike Antkare, One of the Great Stars in the Scientific Firmament, 22th newsletter of the international society for scientometrics and informatrics, 6(2), pp. 48-52.

Marcuzzo M.C. - Zacchia G. (2007), L'ECONLIT e gli strumenti per la valutazione della ricerca economica in Italia, Rivista Italiana degli Economisti, 12 (2), pp. 277-306.

Marcuzzo M.C. - Zacchia G. (2010), Peso alle citazioni o pesi alla numerosità? La valutazione degli economisti accademici italiani, Politica Economica, 26 (3), pp. 409-40.

Meho L.I. - Yang K. (2007), A new era in citation and bibliometric analyses: Web of Science, Scopus, and Google Scholar, in Journal of the American Society for Information Science and Technology, vol. 58 (13), pp. 2105–2125.

OECD (2011), Economic surveys: Italy, OECD.

Perotti R. (2002a), The Italian University System: Rules vs. Incentives. First Conference on Monitoring Italy, ISAE, Rome.

Perotti R. (2002b), Il Bollettino dei concorsi 1,(http://www.igier.unibocconi.it/perotti).

Perotti R. (2003), Il Bollettino dei concorsi 2,(http://www.igier.unibocconi.it/perotti).

Perotti R. (2004), Il Bollettino dei concorsi 3,(http://www.igier.unibocconi.it/perotti).

Perotti R. (2006), Il Bollettino dei concorsi 4,(http://www.igier.unibocconi.it/perotti).

Perotti R. (2008), L'università truccata, Einaudi, Milano.

Schreiber M. (2008), To share the fame in a fair way, hm modifies h for multi-authored manuscripts, New Journal of Physics, vol. 10 (4), pp. 040201.

Tarantino E. (2006), Troppo o troppo poco? Web of science, Scopus, Google Scholar: tre database a confronto (un caso di studio), in Bollettino AIB, vol. 46 (1/2), pp. 23-32.

Woeginger G.J., (2008), An axiomatic analysis of Egghe's g-index, Journal of Infometrics, 2, pp. 364-368.

Wu Q., (2010), The w-index: A significant improvement of the h-index, Journal of the American Society for Information Science and Technology, 61 (3), pp. 609-614.