The role of citation networks to explain academic promotions: an empirical analysis of the Italian national scientific qualification

Maria Cristiana Martini[1] · Elvira Pelle[1] · Francesco Poggi[2] · Andrea Sciandra[1]

Abstract

The aim of this paper is to study the role of citation network measures in the assessment of scientific maturity. Referring to the case of the Italian national scientific qualification (ASN), we investigate if there is a relationship between citation network indices and the results of the researchers' evaluation procedures. In particular, we want to understand if network measures can enhance the prediction accuracy of the results of the evaluation procedures beyond basic performance indices. Moreover, we want to highlight which citation network indices prove to be more relevant in explaining the ASN results, and if quantitative indices used in the citation-based disciplines assessment can replace the citation network measures in non-citation-based disciplines. Data concerning Statistics and Computer Science disciplines are collected from different sources (ASN, Italian Ministry of University and Research, and Scopus) and processed in order to calculate the citation-based measures used in this study. Then, we apply logistic regression models to estimate the effects of network variables. We find that network measures are strongly related to the results of the ASN and significantly improve the explanatory power of the models, especially for the research fields of Statistics. Additionally, citation networks in the specific sub-disciplines are far more relevant than those in the general disciplines. Finally, results show that the citation network measures are not a substitute of the citation-based bibliometric indices.

Keywords Citation networks · Research evaluation · Bibliometrics · Citation-based metrics · ASN · Academic promotion

---

Maria Cristiana Martini, Elvira Pelle, Francesco Poggi, Andrea Sciandra have contributed equally to this work.

✉ Francesco Poggi
   francesco.poggi@cnr.it

Extended author information available on the last page of the article

 🙂 Springer

Introduction

Performance evaluation is a crucial task in every organisation, aimed at monitoring the processes, allocating resources and funds, planning promotions and securing the organisation's effectiveness, efficiency and reputation. Universities and research institutes are no exception, with a double level of evaluation: individuals and institutions. Scholars' and universities' research performance is measured by means of their publication activity, and the results can be used for recruitment decisions, funding allocation, university ranking, and so on.

In Italy, for the assessment of the individual performance, Law 240/2010 (Law, 2011) has introduced a nation-wide research evaluation procedure named National Scientific Qualification (ASN[1]), analogous to the procedures adopted in other European countries, like Germany, France and Spain. In particular, it introduced the ASN as a prerequisite to obtain a permanent position. The evaluation procedure is aimed at certifying if a scholar has reached the scientific maturity expected for the corresponding position of Associate Professor (AP) and Full Professor (FP), and this certification is needed to apply for the position in local evaluation procedures conducted by the Universities. However, passing the ASN does not grant a tenure position. Universities are in charge of creating new positions in compliance with local hiring regulations and on the basis of financial and administrative requirements.

The assessment of the candidates is performed separately for each of the 190 Research Fields (RFs) in which the Italian academic disciplines are segmented; for each RF, a committee composed of five full professors is appointed, and the evaluation is performed considering the CVs submitted by the applicants and three quantitative indices computed for each candidate. Since CVs, indices and the results of the assessments of each candidate are made publicly available, the ASN constitutes an opportunity to perform analyses about a nation-wide evaluation process.

Different scientific fields show different habits both in the publication strategies and in the evaluation tradition. The ASN procedure reflects these differences, and identifies two distinct groups of RFs, improperly named "bibliometric" and "non-bibliometric" fields: the bibliometric fields include STEM and life sciences disciplines; conversely, the humanities and the social sciences are mainly included in the non-bibliometric fields. In fact, although it suggests that only bibliometric disciplines are assessed using bibliographic indices, in the first ASN phase all candidates are assessed using quantitative indices that can be considered in broad sense bibliometric (the main difference being the presence/absence of quantitative measures for the impact of the research activity in the assessment of individual scholars). Therefore, in order to better contextualise this study and its purposes, we decide to rename these disciplines into *Citation-based Disciplines (CDs)* and *Non-citation-based Disciplines (NDs)*. Two different sets of three quantitative indices (one for bibliometric and one for non-bibliometric fields, see Section "Data processing workflow") accompany the qualitative analysis of CVs for the individual evaluation of each applicant. Indeed, two out of the three indicators used to evaluate the first group of disciplines are citation-based; whereas, no citation-based indicators are used to evaluate the second type of disciplines.

---

[1] The acronym ASN stands for *Abilitazione Scientifica Nazionale*. For the rest of the paper, some acronyms (e.g., ASN, ANVUR, etc.) are based on the original Italian names, since they are well established in the Italian scientific community. The English translations are also provided for the benefit of the international readers.

The qualification is not awarded unless at least two of the three indices show a value not lower than some thresholds defined by the National Agency for the Assessment of Universities and Research (ANVUR), a public agency established with the objective of assessing the Italian academic system. These thresholds vary across the RF and the position (FP or AP) for which the candidates apply. For those applicants who awarded at least two thresholds, the evaluation process continues with the assessment performed by the committees on the basis of their CVs, their accomplishments organised into 10 groups (i.e., organisation of or participation in conferences, principal investigator of research projects, membership to journal editorial boards, membership to Ph.D. boards and committees, etc.), and a qualitative assessment of a selection of their publications. Applicants who fail to get a qualification have to wait a year before applying again to the same RF and level. Once acquired, a qualification lasts for nine years.

Several studies have attempted to investigate what are the best predictors of the qualification attainment, and what is the role of bibliometrics in academic promotions and career outcomes (see, e.g., Bedogni et al., 2022; Demetrescu et al., 2020; Jensen et al., 2008; Marini, 2017; Poggi et al., 2019; Tregellas et al., 2018; Vieira et al., 2014), but none of these tried to insert predictors based on citation network indices, as discussed in details in "Related work". As pointed out by Ding (2011), in fact, citation networks contain relevant information to quantify and qualify citations and citation structure, providing a sort of endorsement pattern among the nodes. Moreover, a citation network can be seen as an information network representing the extent to which the research interests of each scholar are related to one another, and can be used to point out the existence of scientific communities (Ji & Jin, 2016). This might be useful when analysing the relationship between bibliometric indices and the attainment of a scientific qualification, since many candidates apply to more than one scientific discipline, often with different results, due to different degrees of consistency between the candidate's research interests and the main contents and scopes of each discipline. The quantitative indices, in fact, do not vary if applied to one RF or another (while the thresholds do), and are then unable to account for the pertinence of the research activity for the discipline, and to help predict the qualitative decisions made by the committee.

The aim of this study is to investigate if the individual position in the citation network can help explain the results of peer review, and if this relationship is due to a correlation between the citation network indices and some of the quantitative indices adopted in the evaluation procedure, or if network data may enhance the model beyond bibliometric ASN indices. In this work we therefore want to investigate the following research questions:

RQ1: Are citation network indices related to the results of the researchers' evaluation procedures?

RQ2: Can network measures enhance the possibility to predict the results of the evaluation procedures beyond the basic performance indices?

Since the evaluation procedure is partly different for CDs and NDs, in that the quantitative indices are not the same, we mean to compare the effect of the citation network measures on the results of the evaluation procedures for the two groups by investigating a supplementary research question:

RQ3: Are citation network measures equally relevant in predicting the procedure's result in the citation-based and in the non-citation-based fields?

Moreover, in case the citation network has a role in the prediction of results, we will try to establish if this role is connected to the bibliometric relevance underlying the citation network structure, or to its capability to mark the borders of a scientific discipline. Then, for each type of discipline (CDs and NDs), we want to analyse a small number of RFs

which overlap to some extent in terms of multiple applications and multiple qualifications, nevertheless maintaining a certain amount of separation. We expect that, if the relevance of the citation network is limited to an indirect bibliometric measure (e.g., proportional to citations count), its importance should be larger in the NDs, that do not have citations count and h-index among the quantitative indices used in the ASN, and fade out if these indices are included in the model. We therefore intend to address a further research question:

RQ4: Can the quantitative indices used in the citation-based disciplines replace the citation network measures in the non-citation-based disciplines?

The study analyses data from the ASN 2016–18, the most recent completely available when we wrote this article; this ASN session spans 5 four-month periods starting from June 29th 2016 to April 6th 2018, and received a total of more than 58,000 applications on the 190 RFs. The majority of candidates only applied to one RF, but a relevant number applied to two or more, as discussed in section "Data processing workflow". This also happened for the first session in 2012–2013, as pointed out by Marzolla (2016).

Among the CDs we choose to focus on the Computer Science, a discipline that has been studied with different purposes both in the Italian context (see, e.g., Demetrescu et al., 2020; Di Iorio et al., 2019) and in the international one (Ding, 2011; Ding et al., 2009). Among NDs we concentrate on Statistics, a discipline that does not have a strong bibliometric tradition, but is going to adopt a citation-based approach in the next Italian assessment for institutions (i.e. the National Research Quality Evaluation—VQR); besides, citation network of statisticians has been analysed by Ji and Jin (2016). Moreover, in spite of the adoption of different evaluation rules, the two disciplines share common habits in terms of coauthorship and citation behaviour, as reported in subsection "Data processing workflow". Last but not least, the authors of this paper are computer scientists and statisticians, and the direct knowledge on the fields may help understand in more depth the results of the analyses.

The remainder of this article is structured as follows: section "Related work" reviews results from previous papers, both on citation networks and on the results of ASN; section "Methods and materials" introduces the data to be analysed and the statistical models to be applied; section "Results" provides the outcomes of the analyses, which are then discussed in the "Discussion" section; finally, "Conclusions and future work" are illustrated in the homonymous section.

Related work

National scientific qualification (ASN)

The ASN constitutes an opportunity to perform analyses about a nation-wide evaluation process. Since the introduction of the ASN as a prerequisite to obtain a permanent position, an increasing number of research papers have been proposed in the literature with the aim to study the impact of the new rules on the Italian university system. Along this direction, Marzolla (2015) provides a first analysis of the ASN procedure through the computation of some quantitative measures and draws attention to potential problems in future appointments of the ASN.

Abramo and D'Angelo (2015) focus on the first wave of the qualification competition with the aim to understand the relationship between scientific merit of the applicant and

the performance to the qualification procedure, including variables that can be indices of possible favoritism and discrimination practices.

The scientific debate also focused on the impact of the new recruitment system on the presence of women in Italian universities (Pautasso, 2015) and on the appropriateness of the quantitative measures (in particular bibliometrics) as indices of the quality and the productivity of scientific research as well as the appropriateness of the feed-back provided to the ASN candidates (Marzolla, 2016).

The introduction of the new regulations of the recruitment system results in an increasing competitiveness. As a consequence, especially for CDs, some authors pointed out some modification in the citation behaviour, in the sense that academics change the citation practice (with an increase of the self-citation phenomenon) with the aim to obtain a value of the citation indices useful for their purposes (see, among others, Peroni et al., 2020; Scarpa et al., 2018). Also the use of alternative indices was investigated (Nuzzolese et al., 2019).

Di Iorio et al. (2019) investigate the performance to ASN from a different point of view, using open data to reproduce the scientific evaluation process, while Poggi et al. (2019) propose a model that allows to predict the performance to ASN on the basis of information in the CV of the applicants and to improve the accuracy of predictions through the identification of a set of quantitative indices.

Demetrescu et al. (2020) focus on CDs (e.g. Computer Science and computer engineering) showing that the use of the same bibliometric measures for all members within the same discipline as thresholds to obtain the ASN, can favour such sub-communities characterised by higher bibliometric indices, and disfavour others.

Citation networks

Information on the scientific merit of academics can be also derived looking at the citation networks in which they are embedded. Indeed, many authors have studied citation networks from several points of view with the aim to understand their main features (Goldberg et al., 2015; Radicchi et al., 2012).

Citing behaviour implies endorsement, confers authority, traces provenance, and it can also be interpreted as scholarly trust (Ding, 2011). Citation practices may be essential to the distribution of symbolic capital and its accumulation by scientists; they can provide insight into the hierarchies within a field and among fields (Wallace et al., 2012).

Many authors investigated the relationship between the citation network and the co-authorship network in order to better understand the mechanism of scientific collaboration among academics of different institutions. Indeed, it is well-known that scientific collaboration plays a key role in enhancing scholarly communication. On the other hand, citing behaviour may be a sign of acknowledgement of the research work, conferring authority to the cited author(s). In particular, Ding (2011) states that there is a direct co-author relationship between productive authors with the same research interests, while they generally do not collaborate when they do not share research issues. Conversely, highly cited authors do not generally coauthor with each other, but closely cite each other.

Wallace et al. (2012) use the co-authorship network in order to explain the proximity of authors in the citation network, while Ji and Jin (2016) studied the relationship between co-authorship and citation networks among statisticians.

A different approach to the study of citation networks is represented by the use of PageRank algorithms. In this regard, Ding et al. (2009) propose a weighted PageRank

algorithm to rank authors, and Singh et al. (2011) discuss another version of the PageRank algorithm developed to rank research papers and, consequently, assign scores to both conferences and authors in order to rank them.

Methods and materials

This section introduces all the methods and material used for our study. The data gathered and the software developed for our work are available in (Martini et al., 2021) and on GitHub at https://github.com/DigitalDataLab/ASN16-18_CitationNetwork.

Data collection: input data

For the purposes of this study we focused on two disciplines, one citation-based, i.e. Computer Science, and the other non-citation-based, i.e. Statistics. In the Italian academic system, Computer Science is organised in two RFs: Informatics (code 01/B1) and Information Processing Systems (code 09/H1). Statistics is organised in three RFs: (Methodological) Statistics (code 13/D1), Economic Statistics (code 13/D2) and Demography and Social Statistics (code 13/D3).

In order to investigate the relation between citation networks and the results of researchers' evaluation procedures, we need to make up the citation structure among scientists belonging to the same RF; here, given the national nature of the ASN procedure, we limited our analyses to the part of citation networks that involves Italian authors. Then, for each of the five aforementioned RFs we considered (a) the candidates to the ASN and (b) the permanent positions (i.e. FPs, APs and Assistant Professors) within the Italian academic system.

We focused on the candidates to the 2016–2018 session of the ASN, as the following 2018–2021 session was still in progress at the time of writing this article. We collected 2950 candidate's CVs in PDF from the ANVUR website,[2] where they were made available for a short period of time. Table 1 reports the number of candidates to the five sub-sessions of the 2016–2018 ASN and the applicants' success rate (point a).

Table 2 reports the permanent positions in the considered RFs in the Italian university system as of 31/12/2016, i.e. at the beginning of the 2016–2018 ASN session (point b). The list of 2324 permanent positions has been downloaded from CercaUniversita,[3] a website containing information about Italian academics. In particular, for each person we collected their name, surname, gender, university, department, role and RF.

The candidates' CVs and the list of permanent positions are the input of the process described in the following section.

Data processing workflow

In order to calculate the citation-based measures used in this study, we needed to collect the bibliographic and citation data of the Italian academics, i.e. the ASN candidates and the permanent positions within the Italian university system, in the five considered RFs. To

---

[2] https://abilitazione.miur.it/.

[3] https://cercauniversita.cineca.it/.

Table 1 Applications (i.e. CVs) submitted to the five sub-sessions (S1–S5) of the 2016–2018 ASN in Computer Science disciplines (i.e. RFs 01/B1 and 09/H1) and Statistics disciplines (i.e. RFs 13/D1, 13/D2 and 13/D3). We reported the number of applications and the success rate (in parentheses)

Discipline	Recruitment Field (RF)	Recruitment Field (RF) Code	Associate Professor (AP)						Full Professor (FP)						AP + FP
			S1	S2	S3	S4	S5	Total	S1	S2	S3	S4	S5	Total	S1-S5
Computer Science	Informatics	01/B1	307 (30.6%)	58 (25.9%)	45 (46.7%)	88 (40.9%)	259 (39.0%)	**757 (35.3%)**	235 (31.9%)	34 (32.4%)	22 (40.9%)	83 (47.0%)	144 (41.7%)	**518 (37.5%)**	**1275 (36.2%)**
	Information Processing Systems	09/H1	271 (41.7%)	51 (31.4%)	38 (50.0%)	83 (45.8%)	228 (50.4%)	**671 (44.9%)**	171 (36.8%)	23 (47.8%)	26 (61.5%)	61 (52.5%)	132 (59.1%)	**413 (48.4%)**	**1084 (46.2%)**
TOTAL - COMPUTER SCIENCE			578 (35.8%)	109 (28.4%)	83 48.2(%)	171 43.4(%)	487 (44.4%)	**1428 (39.8%)**	406 (34.0%)	57 (38.6%)	48 (52.1%)	144 (49.3%)	276 (50.0%)	**931 (42.3%)**	**2359 (40.8%)**
Statistics	(Methodological) Statistics	13/D1	64 (50.0%)	15 (60.0%)	4 (50.0)	18 (72.2%)	45 (31.1%)	**146 (47.9%)**	50 (74.0%)	9 (66.7%)	4 (50.0%)	13 (38.5%)	33 (39.4%)	**109 (57.8%)**	**255 (52.2%)**
	Economic Statistics	13/D2	46 (41.3%)	7 (28.8%)	7 (28.6%)	8 (12.5%)	49 (32.7%)	**117 (34.2%)**	28 (50.0%)	5 (40.0%)	5 (60.0%)	5 (40.0%)	25 (28.0%)	**68 (41.2%)**	**185 (36.8%)**
	Demography and Social Statistics	13/D3	39 (35.9%)	8 (25.0%)	4 (50.0%)	12 (16.7%)	43 (48.9%)	**106 (38.7%)**	19 (42.1%)	3 (33.3%)	3 (0.0%)	2 (50.0%)	18 (61.1%)	**45 (46.7%)**	**151 (41.1%)**
TOTAL - STATISTICS			149 (43.7%)	30 (43.4%)	15 (40.0%)	38 (42.1%)	137 (37.2%)	**369 (40.9%)**	97 (60.8%)	17 (52.9%)	12 (41.7%)	20 (40.0%)	76 (40.8%)	**222 (50.5%)**	**591 (44.5%)**

Table 2 Permanent positions (i.e. FPs, APs and Assistant Professors) in Computer Science disciplines (i.e. RFs 01/B1 and 09/H1) and Statistics disciplines (i.e. RFs 13/D1, 13/D2 and 13/D3) in 2016 within the Italian academic system

			Assistant Professor	Associate Professor (AP)	Full Professor (FP)	Total
Computer Science	Informatics	01/B1	362	312	192	**866**
	Information Processing Systems	09/H1	261	299	182	**742**
TOTAL - COMPUTER SCIENCE			**623**	**611**	**374**	**1608**
Statistics	(Methodological) Statistics	13/D1	159	170	111	**440**
	Economic Statistics	13/D2	49	54	41	**144**
	Demography and Social Statistics	13/D3	45	53	34	**132**
TOTAL - STATISTICS			**253**	**277**	**186**	**716**

achieve this result, we used the four step procedure shown in Fig. 1 and described below. In this process we decided to use Scopus to collect data about authors and their publications,
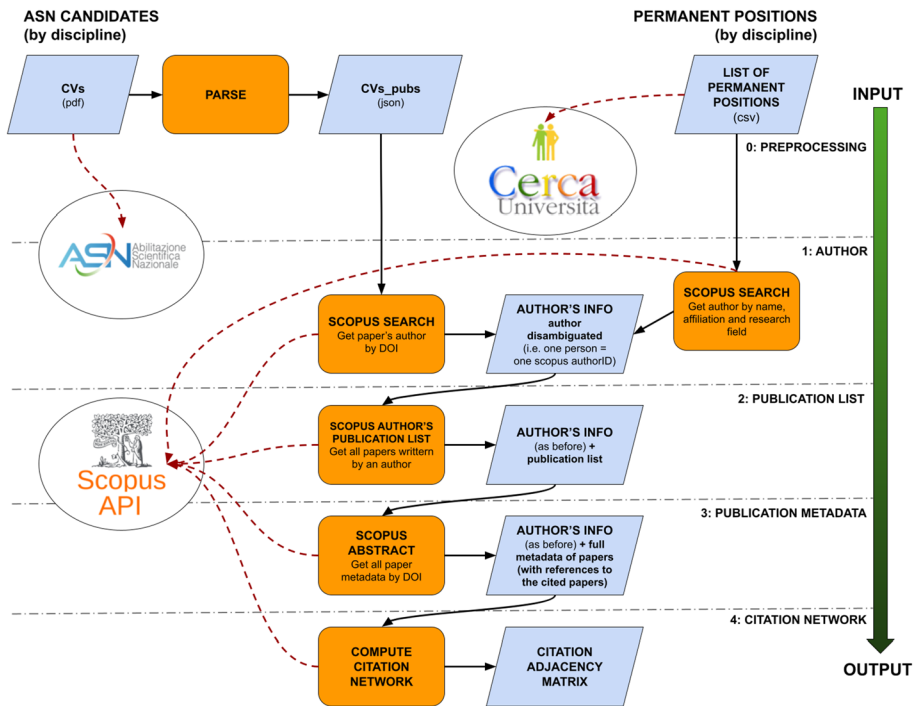
Fig. 1 Data processing stages (orange rectangles) with external data sources (white circles) and outcomes (blue parallelograms) for each discipline. On the top-left side candidates CVs are ingested in PDF format, while on the top-right side the list of permanent positions is collected in CSV format. The citation matrix summarising the citation network of a discipline is computed in four consecutive stages, which are depicted from the top to the bottom

as it is a very extensive and reliable bibliometric database, and because it is the official data source used by ANVUR to calculate the indicators used in the ASN.

The objective of the first step is to disambiguate each person (i.e. the academics who submitted an application to the ASN and those with a permanent position), and to get the corresponding Scopus author IDentifier (authorID) that uniquely identifies each author in Scopus. The authorID is used in the following steps to retrieve all authors information (e.g. publication list, paper citations, etc.) needed to compute the citation networks and the indices used in this study from Elsevier's Scopus database. To access this data we used the Scopus API,[4] the official interface to Scopus services exposed by Elsevier.

To retrieve applicants' authorIDs, we started by extracting the DOIs of each publication listed in the candidates' resumes. The resumes are in PDF format, and were downloaded from the ANVUR website where they have been made publicly available for a short period of time. We converted the resumes from PDF into a pure textual format (i.e. TXT), then we used regular expressions to extract structured information (e.g. the title, authors and DOIs of the publications), and stored it into JSON files, as depicted in Fig. 2. For each ASN applicant we used the DOIs of their publications to retrieve basic metadata for each paper,
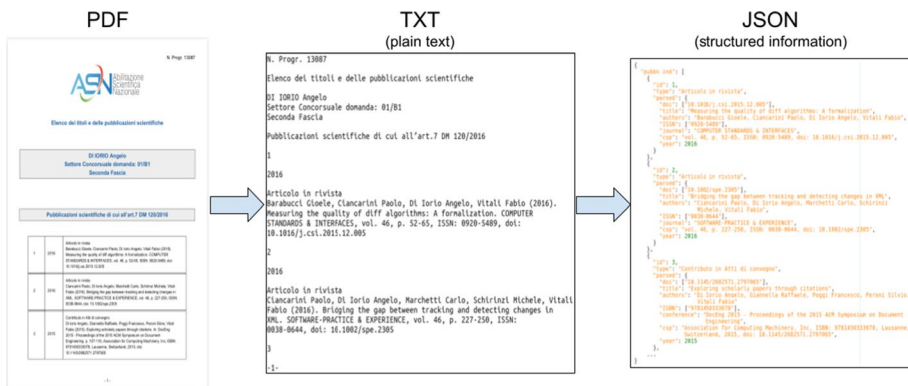

Fig. 2 An example of the process used to convert an applicant's CV from PDF to TXT (i.e. plain text), and then to a JSON file containing structured information

including authors information. The identification of the author who is common to all the publications in a CV allows us to find out the Scopus authorID of the CV's titleholder.

To get the authorIDs of the academics with a permanent position we used a different strategy. Since we do not have the list of their publications, we searched authors on Scopus using their names, affiliations and research fields as keys in the query. In cases where the search has returned more than one identifier we performed a manual check.

At the end of this stage we successfully disambiguated 3,947 out of 3,967 ASN applications and permanent positions (99.50% of the total) in Computer Science, and 1253 out of 1307 (95.94% of the total) in Statistics, as reported in Table 3. We manually checked the remaining 74 cases and observed that they were mostly empty or incomplete ASN applications (i.e. the publication list was missing in the CV) and young academics with permanent positions who are not in Scopus, so we excluded them from our analysis.

It is important to note that, although the ASN is bound to a specific RF and professional level, it is possible to apply in different RFs and roles. For example, in the 2016–2018 session, in Computer Science 427/588 (76.6%) applicants for AP in the RF 09/H1 (Information Processing Systems) also applied to 01/B1 (Informatics). In the same session, out of 331 academics who applied to the ASN for AP in Statistics disciplines (i.e. RFs 13/D1, 13/D2 and 13/D3), 58 (17.5%) submitted their applications in at least two RFs, and 15 (4.5%) in all three RFs. As a consequence, a Scoups authorID can be associated with multiple ASN applications and/or permanent positions. For instance, authorID 23,398,635,500 is associated with an application in the RF 09/H1 at the AP level (first sub-session), two applications in the RF 01/B1 at the AP level (first and fifth sub-sessions), and a permanent position (i.e. Assistant Professors) in the RF 09/H1. In total, our dataset is composed of two groups of academics uniquely identified by an authorID: one is composed of 2,145 academics in Computer Science (RFs 01/B1 and 09/H1), the other of 889 academics in Statistics (RFs 13/D1, 13/D2, 13/D3).

In the next stage we queried the Scopus database to retrieve the list of publications of each academic. We retrieved 112,599 publications for the 2145 disambiguated academics in the Computer Science discipline, and 17,752 publications for the 889 disambiguated academics in the Statistics discipline.

Table 3 Results of the disambiguation stage. The table reports the number of ASN applications and permanent positions to which a Scopus authorID has/has not been associated in Computer Science and Statistics

	Computer Science			Statistics		
	Disamb.:success	Disamb.:failure	Total	Disamb.:success	Disamb.:failure	Total
ASN applications (CVs)	2345	14	2359	568	23	591
Permanent positions	1602	6	1608	685	31	716
Total	**3947 (99.5%)**	**20 (0.5%)**	**3967 (100.0%)**	**1253 (95.9%)**	**54 (4.1%)**	**1307 (100.0%)**

Table 4 Information about the two computed citation networks

	Computer Science	Statistics
Academics (i.e. authorIDs)	2,145	889
Papers	112,599	17,752
Average of publications authored by each academic[a]	85.44	25.16
Average authors per paper	3.62	3.48
Median of authors/paper	3	3
Average citations per paper	25.61	36.14
Median of citations/paper	21	31
Citations (links) inside the network	293,422 (10.93% of total citations)	24,272 (3.96% of total citations)
Citations inside the network/paper	2.61	1.38

[a]Since a paper may be written by two or more authors, this average is higher than the number of publications divided by the number of academics in the discipline.

Then, we queried Scopus to collect detailed metadata for each of the retrieved publications. This metadata contains the information needed to compute the citation network and the other indices used in this study.

Finally, we used the reference list in the metadata of the papers to compute two main citation networks: one for the 2145 Italian computer scientists (i.e. RFs 01/B1 and 09/H1), and the other for the 889 Italian statisticians (i.e. RFs 13/D1, 13/D2 and 13/D3). The citation network of the computer scientists contains 293,422 links (i.e. DOI-to-DOI citations within the network), and that of the statisticians contains 24,276 links. Table 4 summarises some basic information about these citation networks. Moreover, we also computed five sub-networks, i.e. one for each of the five RFs under investigation, as described in Section "Results".

By inspecting the table we note that on average computer scientists write more articles than statisticians, while the number of authors for each publication is very similar. We also observe a difference concerning the citations: statisticians insert in their papers more citations than computer scientists, but very few fall into the network of Italian statisticians. On the contrary, computer scientists insert fewer citations, but the percentage of those that fall within the network of their discipline is much higher than that of statisticians.

In the event of non-qualification, candidates may reapply for the same RF and the same position (AP or FP) one year later. In this way, multiple applications of the same researcher for the same position and RF are included in the available data. This results in autocorrelated datasets with non-independent observations. To avoid duplication of the same statistical unit in multiple applications for the same RF and position, we chose to focus on what happened at the end of the whole ASN session (qualification or non-qualification in each RF and position). Indeed, we knew for each candidate whether or not they succeeded in the qualifications in which they participated in the whole session. Consequently, we decided to use for each candidate only the last result achieved in the ASN session, and to collect the citation networks data at the end of the session (2018-04-07).

From the data collected in the process we extracted the variables used in the models developed for all the analyses presented in this work. In particular, for each ASN candidate in Computer Science and Statistics, we focused on three group of variables:

– The *ASN scientific indices:* the values of the indices computed for each applicant by the ANVUR and published on the ASN website. The data necessary for this computation are retrieved from Scopus and Web of Science, and only publications that are less than 15 years old are taken into consideration for candidates to the role of FP, while only those that are less than 10 years old are considered for candidates to the role of AP. In particular, we consider the following three indices for Computer Science:

- CD_I1: the number of their journal papers;
- CD_I2: the total number of citations received;
- CD_I3: their h-index;

and the following three indices for Statistics:

- ND_I1: number of their journal papers and book chapters;
- ND_I2: number of their papers published on top-class journals[5];
- ND_I3: number of their published books.

– The *ASN accomplishments*: the list of accomplishments awarded to each candidate by the committee has been extracted from the final reports published on the ASN website;
– The *citation network measures*: these measures are described in Section "Citation network measures", and have been computed using the citation matrices based on the information about citations collected in the aforementioned five-stage process.

Citation network measures

To create the citation network of the selected authors, we defined the edges according to the citations of their papers (Dawson et al., 2014). By viewing citation data as a network, we can use different measures to extract useful information. Graph-level indices provide a summary of structural properties pertaining to the global structure, e.g.:

- *Size* The number of vertices.
- *Edge count* The number of edges within a graph.
- *Density* The ratio of the number of edges and the number of potential edges.
- *Reciprocity* The proportion of dyads that are symmetric (i.e., mutual connections in a directed graph).
- *Centralization* The absolute deviation from the maximum value of a centrality index (in this work we chose the indegree centrality) for a graph (Freeman, 1978).
- *Hierarchy* The extent of asymmetry in a graph. We measured the hierarchy level of a graph by Krackhardt (1994) hierarchy score, computed as the proportion of asymmetrical dyads in the reachability graph.

We compared (see Section "Results") the structure of the global networks of Italian academics of Computer Science and Statistics by selecting some graph-level measures (density, reciprocity, centralization, and hierarchy) that may suggest the existence of different traits at a descriptive level.

---

5 The list of top-class journals is released by ANVUR and is available at https://www.anvur.it/attivita/class ificazione-delle-riviste/classificazione-delle-riviste-ai-fini-dellabilitazione-scientifica-nazionale/elenchi-di-riviste-scientifiche-e-di-classe-a/ (Last accessed 11 May 2021).

Instead, in the context of citation networks, the focus on the local properties of a graph generally involves measures based on the number of incoming citations of an author (node-level indices). A simple citation count can quantify the success of a scientist, because the number of citations obtained by a paper are then transferred to its authors to assess their scientific quality (Radicchi et al., 2012). Examining the citation process in a network frame allowed us to focus on the local properties of the citation network. For instance, centrality is a key concept in order to identify the actors that are structurally more central than the other nodes in the network (Diallo et al., 2016). Since a citation network is a directed graph, we did not consider measures based on geodesic distances, such as closeness and betweenness (Freeman, 1978). Instead, an important reference is Freeman's degree centrality measure (Freeman, 1978), which counts the direct ties each node possesses within the network, in order to quantify a researcher's impact (also on a specific field). The degree centrality works through the direct ties of a node, with the aim of capturing its connectedness or popularity (Liu et al., 2015). In particular, we chose to focus on the following centrality measures:

- *Indegree*, i.e., the number of scientists who cited a scientist. Therefore, indegree is a fundamental measure counting how many times a scientist has been cited within a citation network. Generally, nodes with a higher indegree are structurally more central and, consequently, show a higher ability to influence others (Yan & Ding, 2009).
- *Outdegree*, i.e., the number of scientists cited by a scientist. In a citation network, outdegree consists of the scientists that inspired a scientist's work, providing a measure of how immediately influenced (Borgatti, 2005) is a scientist.

As the degree of a node is given as a function of its adjacent nodes, it can be interpreted as a measure of local centrality, our main target since we are studying a citation network. However, we are also interested in testing other local measures to capture different node's properties:

- *Eigenvector Centrality (evcent)* is calculated on the basis of the principal eigenvector of the adjacency matrix defining a network (Bonacich, 1972). The idea beyond the evcent score is that a node is more central if it is connected to nodes that are themselves central. Evcent was used to measure position centrality in a faculty hiring network (Feeley et al., 2011), a similar scope to the ASN. Another relevant feature is that it can take into account multiple relationships between two network nodes. We expect authors with high evcent to be more central than other authors with lower values, because an increase in the number of citations for an author in the network determines an increase in the value of the eigenvector for all the cited authors (Diallo et al., 2016).
- *Bonacich's Power (bonpow)* calculates the influence of a node recursively, commonly by weighting more the nodes connected to other nodes that in turn show high influence (Liu et al., 2015). In particular, the score distribution is controlled by an exponential parameter which determines the type and the intensity of nodes' dependency upon their *alters*: when the exponent is equal to 0, the score is proportional to the outdegree and it is equivalent to evcent when the exponent is set to the reciprocal of the first eigenvalue of the adjacency matrix (Butts, 2008). When the exponent is positive, *ego* becomes stronger when connected to strong alters and vice versa. A small magnitude of the parameter implies a heavy weight of the local structure, while a large magnitude considers more the position of a node in the whole structure (Bonacich, 1987).

Therefore, we set this parameter at a positive small value (0.05), as an author's power is increased by being cited by other prominent scientists.

- *A Prestige Index*, in this context, measures the extent to which scientists acquire direct citations by other scientists of the same field. Among the range of prestige indices, we chose the indegree within the row-normalised graph (Butts, 2008; Wasserman & Faust, 1994): for a given scientist, it measures the proportion of citations received from any other scientist of the field, divided by their total citations (i.e., how much a cited scientist is important in each citing bibliography).

- *The Information Centrality (infocent)* of a node is calculated by considering every possible path between node pairs, through the harmonic mean of the node's distance to the others (Stephenson & Zelen, 1989). Infocent measures the extent by which each node has a high number of short paths to other nodes, weighting short paths more heavily than the evcent (Butts, 2008). We expect this index to have a positive relationship with the achievement of the qualification.

- A further measure that we considered is the *Number of Self-Citations* of each scientist, an index that could play a remedial role with respect to other measures. Indeed, in Computer Science RF, the CD_I2 and CD_I3 indices are not adjusted for self-citations, and the use of this measure in our models is intended to take this feature into account.

Models

RQ1 aims at testing whether citation networks indices are related to the results of researchers' assessment procedures. The approach to answer this question is to compare the ASN successful and unsuccessful candidates through citation networks indices. Since the network indices do not follow a normal distribution (See "Normality Assessment of Citation Network Indices" in Martini et al., 2021), we chose the Wilcoxon–Mann–Whitney test applying the continuity correction in the normal approximation for the *p*-value.

To answer RQ2–4, we considered some models in order to explain the ASN results and to estimate the effects of network variables in the Computer Science and Statistics RFs. Since the ASN result is a binary response (two classes), we needed to choose a method for classification. Among several available alternatives (Linear Discriminant Analysis, Classification Trees, etc.) we chose linear logistic regression. Our goal is not to maximise the classification accuracy of the ASN results, but rather to achieve an interpretable description of how the independent variables affect the dependent variable. Specifically, logistic regression models would allow us to understand the role of the network variables in explaining the ASN results.

We selected the predictors by the forward stepwise selection. The forward stepwise selection builds a model sequentially, adding one variable at each step, by identifying the predictor that most improves the fit and by including it in the model. The forward stepwise selection procedure tries to maximise an optimality criterion function, through a sequential *F* test based on a fixed level (in our case 5%).[6] Specifically, this selection method maximises the squared partial correlation coefficient with the dependent variable, given the set

---

[6] We also tested the Akaike information criterion (AIC) and the Bayesian information criterion (BIC) selection procedures, which showed essentially identical results. The main difference found is that BIC showed a convergence problem for one sub-sector, so we will only comment on the results achieved on the basis of the significance of the *F*-statistic.

of variables already selected (Bendel & Afifi, 1977). In this way, the procedure selects an independent variable by choosing the one having the maximum *F* value as derived from the test of the hypothesis (null hypothesis: the associated partial correlation coefficient is equal to zero). Given the small size of most of the investigated fields, we did not randomly split the data into training and test sets, because it would not be a reliable procedure due to a probable overfitting.

To assess the RQ2–4 statements we will use: the pseudo R-squared (McFadden) of the models of each field; the *p*-values for the predictors' significance; the values (linear log-odds) and the signs of the significant predictors' parameter estimates (e.g.: with a positive coefficient, an increase of 1 citation accounts for an increase in the odds of the ASN attainment; a negative coefficient for a predictor suggests that, all other variables being equal, the ASN attainment is less likely to occur; etc.).

Although the forward stepwise logistic regression may lead to biassed estimates, it will have lower variance and we will make less assumptions than other methods, e.g. Linear Discriminant Analysis (Hastie et al., 2001). Even if it does not guarantee to find the best possible model, the forward stepwise method fits the data in a search for a parsimonious model. In this way, this method involves a subset of the features, making the models more interpretable for our purposes. We also chose the forward stepwise method in favour of clarity, not least because it reproduces the evaluation process (by steps) of the ASN.

Operational decisions

Concerning RQ2 and RQ3, our analysis strategy is based on the use of logistic regressions with stepwise-forward selection for all sub-fields of Computer Science and Statistics. The variables included in the models can be grouped into the following blocks: ASN scientific indices, accomplishments, and network measures.

The ASN indices CD_I1-CD_I3 (citation based disciplines) and ND_I1-ND_I3 (non citation based disciplines) are meant to measure the scientific performances of the candidates for the ASN (see Section "Data processing workflow").

Another mandatory feature of ASN deals with the number of other positively assessed scientific accomplishments. Each ASN committee identifies up to 10 accomplishments, and each candidate must be positively assessed on at least three accomplishments to be eligible for qualification. For Informatics (01/B1) RF, the number of positively assessed accomplishments was not available, so it had been replaced by the number of accomplishments submitted by candidates under each heading.

The indices we computed for the citation network graphs (Section "Citation network measures") were separated into two groups: the first one with the indices obtained considering the global network graph of each discipline (Computer Science or Statistics), the second one considering each network graph of the RFs. In this way we meant to assess and compare the importance of the candidates' position both at the global level of their discipline and within their own field. Moreover, comparing these different networks in terms of their explanatory capability could allow us to observe whether scientific RFs tend to mark their boundaries in candidates' assessments.

While analysing the distributions of the network measures, we noticed that in the Statistical RFs the infocent index shows only 5 levels, all of which are very close to 0 (mean $= 1{,}02E{-}11$; St.Dev. $= 3{,}54E{-}12$). This result might be due to a very low level of centrality of the members of the Statistics citation network. For the Computer Science RFs we did not detect any anomaly for the infocent distribution. We therefore decided

Table 5 Global citation network measures for the whole disciplines and the recruitment fields

Measure	Computer Science	Statistics	01/B1	09/H1	13/D1	13/D2	13/D3
Size	2145	889	1488	1337	520	261	221
Edge count	89,393	7526	41,939	42,196	4490	884	1042
Density	0.019	0.010	0.019	0.024	0.017	0.013	0.021
Reciprocity	0.978	0.989	0.979	0.974	0.983	0.988	0.978
Centralization	0.132	0.080	0.089	0.150	0.105	0.094	0.146
Hierarchy	0.040	0.304	0.044	0.037	0.253	0.531	0.484

not to consider the infocent index in the multivariate analyses of the Statistics fields. Moreover, by looking at the correlations among the indices, we noticed very large and significant values between outdegree and evcent and between indegree and prestige. In particular, with reference to the complete network graph for Computer Science, Pearson's correlation coefficient between outdegree and evcent is equal to 0.93 ($p < 0.001$), while the one between indegree and prestige is equal to 0.91 ($p < 0.001$). With reference to the complete network of Statistics, the correlation between outdegree and evcent is equal to 0.91 ($p < 0.001$) and that between indegree and prestige is equal to 0.88 ($p < 0.001$). Clearly these measures could be redundant if used together in the multivariate analysis, therefore we made a selection. We decided not to consider the evcent and to choose the outdegree because at a theoretical level it reflects a concept we want to observe (to which extent a scientist is immediately influenced by peers in the same discipline). Instead, we decided to keep both the indegree and the prestige because, although defined on the same basis (citation count), they express different concepts, i.e., the number of direct citations and the importance of a scientist in the bibliography of others in the same field.

The process of variable selection by forward stepwise involves adding one block at a time, as follows:

1. the ASN indices were all entered in the first step (no selection).
2. the scientific accomplishments were selected by the forward stepwise method with entry testing based on the significance of the score statistic ($p < 0.05$) to improve the fit of the model.
3. the global and RFs network measures were also selected by the forward stepwise method.

To answer RQ4, for the Statistical fields we first tried to add all three indices equivalent to those of the CDs to the first block (without any selection), and then we proceeded in the same way described above for the subsequent blocks.

Table 6 Wilcoxon–Mann–Whitney tests for the citation network indices between the ASN successful and unsuccessful candidates groups. ***Denotes significance at the 0.1% level. Indices are computed using the entire Computer Science and Statistics networks, respectively

Index	Successful candidates mean	Unsuccessful candidates mean	W	p
Indegree	36.60	22.24	550,124	0.000***
Outdegree	39.96	27.18	565,101	0.000***
Bonpow	− 0.04	− 0.10	786,559	0.731
Prestige	1.00	0.74	580,875	0.000***
Self citations	126.42	89.60	652,340	0.000***

Results

Comparison between the global networks of the two disciplines and RFs

As explained in section "Citation network measures", the first step of our analysis is to compare Computer Science and Statistics fields in terms of citation networks at graph level. In particular, we exploited some global measures computed for the whole citation networks of Italian academics in the disciplines of Computer Science and Statistics, and for the recruitment fields of Informatics (01/B1), Information Processing Systems (09/H1), (Methodological) Statistics (13/D1), Economic Statistics (13/D2), Demography & Social Statistics (13/D3). Table 5 shows the following graph-level indices (Freeman, 1978; Wasserman & Faust, 1994): node count (size), edge count, density, reciprocity, centralization (based on indegree centrality score), and hierarchy.

Graph-level indices suggest that:

- Computer Science fields are larger in terms of size and, more than proportionally, also in terms of edges since they have higher density.
- Reciprocity turned out to be very high in every field, i.e., we found a high level of mutual citations between authors.
- The dependency on the central nodes of each network (centralization based on indegree centrality) shows rather low levels in all the fields, overall, with higher values in the Computer Science fields than in the Statistics fields.
- One of the most interesting results is the value of Hierarchy of the Statistics discipline, very high compared to the Computer Science one (almost 8 times for the whole field—0.30 vs 0.04—with the highest values in 13/D2 and 13/D3 fields). This may point at the existence of some recognized "schools" in this field. A possible cause of this phenomenon may lie in the citation structure of the three Statistics recruitment fields that appear rather heterogeneous for the topics dealt with.

Citation network indices assessment

For the purpose of answering RQ1, we compared the citation network indices (node-graph level) between the group of ASN successful candidates and the group of ASN unsuccessful candidates. The analysis was carried out on all Computer Science and Statistics candidates, regardless of discipline, because the aim of this step is to assess whether the network indices differ significantly between the two groups, showing

Table 7 Final results of logistic regressions for each Computer Science position and RF. ***, **, *, and °Denote significance at the 0.1%, 1%, 5%, and 10% levels

	Parameter estimates			
	01/B1 AP	01/B1 FP	09/H1 AP	09/H1 FP
(Intercept)	− 5.008***	− 5.454***	− 7.898***	− 4.868***
CD_I1	0.006	− 0.0139	0.009	− 0.013
CD_I2	− 0.001*	− 0.001*	− 0.001°	0.000
CD_I3	0.209***	0.260***	0.161*	0.070
Accomplishments			1.040***	1.114***
Accomplishment 3		0.062*		
Accomplishment 6	0.207***	0.093***		
Accomplishment 9	0.112*	0.107*		
Infocent			0.287***	
RF_indegree	0.031*	0.048***		
RF_infocent	0.177**			
RF_outdegree			− 0.026*	0.022***
Mc Fadden Pseudo R^2	0.210	0.244	0.318	0.248
Accuracy	0.7062	0.747	0.8076	0.738

them to be a priori potential explanatory factors for the outcome of the ASN. Table 6 shows the results of Wilcoxon–Mann–Whitney test performed on each node-level index described in section "Citation network measures". Infocent was not tested because its distribution in the RFs of Statistics showed few levels tending to 0. All the indices tested were significantly different in the two groups, with averages always higher in the successful candidates' group. The only index that is not significantly different is bonpow (Bonacich's power). These results support our choice to use indices obtained from citation networks as explanatory variables of ASN.

In the following sections we present the final results of the stepwise logistic regression models computed for each level (i.e. Full Professor and Associate Professor) and Recruitment Field in the Computer Science and Statistics disciplines. The detailed results on each step of the stepwise procedures are in Martini et al., (2021).

Computer Science

Computer Science logistic regression models show accuracy levels ranging from 70.6% to 80.8% and pseudo R^2 ranging from 21.0 to 31.8%. All final models include citation network indices that play a significant role in explaining the ASN outcome (Table 7).

ASN scientific indices Model results of the first block, consisting of the three ASN indices CD_I1-CD_I3, show that the CD_I1 index (number of articles on Scopus or Web of Science) is never significant in the RFs of Informatics (01/B1) and Information Processing Systems (09/H1). On the other hand, the CD_I2 index (total number of citations received) is significant only in the AP role of both fields, while the CD_I3 index (h-index) is significant in all models and with the lowest *p*-values in each RF. Globally, only the h-index turns out to be a good predictor of the ASN outcome for Computer Science. In fact, the number

of citations shows a negative coefficient (very close to 0) when it is significant. Since both indices CD_I2 and CD_I3 are based on the number of citations, this could mean that only the h-index manages to capture efficiently part of the variability of the ASN outcome. The initial accuracy of the Computer Science models with only the ASN indices is rather limited in each RF (62%-63%) as well as the pseudo R^2 (0.04–0.10).

In summary, *the h-index proves to be the best predictor of the Computer Science applications outcome among the ASN indices*.

ASN accomplishments Adding to the models the candidates' accomplishments, in the fields of Information Processing Systems we find that the number of positively assessed accomplishments is significant and with a positive influence on the dependent variable. The accuracy of these models improves by about 10% and the pseudo R^2 also increases by about 0.1.

Conversely, for the RF of Informatics, the number of assessed accomplishments was not available, so we used the number of items submitted by the candidates for each accomplishment as proxy variables. The selection process reveals two significant accomplishments: number 6 (membership of a doctoral school faculty) for APs and FPs, and number 9 (achievements in technology transfer) for FPs. In this RF, the improvement in terms of accuracy (about $+6\%$) and pseudo R^2 (about $+0.6$) is lower than in the Information Processing Systems RF, most likely due to the absence of the more accurate index on the accomplishments.

Overall, *the number of positively assessed accomplishments, when available, significantly improves the accuracy of the Computer Science models*.

Citation network measures The addition of the third block, consisting of the citation network indices (computed both at discipline and RF level), display the significance of some variables: RF outdegree (09/H1—FP and AP), Computer Science infocent (09/H1—AP), RF indegree (01/B1—FP and AP), and RF infocent (01/B1—AP). The improvement in terms of pseudo R^2 is about 0.09 for Informatics, while it is lower for Information Processing Systems, probably due to the inclusion of the accomplishments index which is one of the requirements for a positive outcome of the assessment procedure. The accuracy improves for all models, in particular for AP in RF 09/H1 ($+6\%$) where it reaches a score of 80.8%.

In the RFs of Computer Science, therefore, some citation network indices are significant and help to explain the result of the ASN, although their impact is quite limited in terms of accuracy and pseudo R^2. It is likely that this result is due to the presence of citation indices such as the h-index, which could limit the explanatory power of the network measures.

Interestingly, almost all of the citation network indices refer to the RF network. This result emphasises the relevance of the candidates' centrality in their field networks, in terms of both received and done citations.

To sum up, *RF network indices result to be significant in the Computer Science models and contribute to improve accuracy, albeit to a limited extent*.

Statistics

ASN scientific indices Logistic regression models fitted to explain the probability of success in the ASN procedure by means of the three quantitative indices used in the evaluation process, in general have a very low accuracy, ranging from 53.2 to 67.6%; also the pseudo R^2 does not reach good levels (see Table 8). In fact, in five out of the six basic models none

Table 8 Final results of logistic regressions for each Statistics position and RF. ***, **, *, and °Denote significance at the 0.1%, 1%, 5%, and 10% levels

	Parameter estimates					
	13/D1 AP	13/D1 FP	13/D2 AP	13/D2 FP	13/D3 AP	13/D3 FP
(Intercept)	− 1.631*	− 2.081**	− 1.675**	− 1.313*	− 3.562**	− 1.471
ND_I1	0.022	− 0.006	− 0.021	0.012	− 0.021	0.028
ND_I2	0.111	0.164**	0.067	− 0.039	0.254°	− 0.111
ND_I3	0.690	0.039	− 0.159	− 0.154	0.594°	− 0.160
Accomplishments					0.366°	
Bonpow	0.506*	0.506*				
Indegree	− 0.488**					
Self citations	− 0.039**				− 0.095*	
RF_indegree	0.726***	0.071*		0.369***		
RF_outdegree	0.138*		0.541***		0.563***	0.306**
Mc Fadden Pseudo R^2	0.275	0.190	0.187	0.181	0.371	0.221
Accuracy	0.7209	0.7315	0.7358	0.6935	0.7778	0.8372

of the indices have a significant effect at a 5% significance level (and not even at the 10%); only for full professorship in Methodological Statistics one of the quantitative indices, the number or journal articles published in a list of top-class journals, have a significant positive effect on the success probability in the ASN procedure.

In short, *models based on the mere ASN scientific indices cannot explain well the probability of success*.

ASN accomplishments Adding to the model the number of accomplishments positively evaluated by the committee does not change the model performance very much, since this characteristic is selected only in one of the models (namely, the one for the associate professorship in Demography and Social Statistics); as predictable, the number of positively evaluated accomplishments is positively associated to the probability of success in the ASN procedure.

ASN accomplishments do not add relevant information.

Citation network measures Introducing the variables that describe the individual's position in the citation network, models' accuracy improves to levels ranging from 69.4 to 83.7%. We can observe that positions in the RF disciplinary citation network are far more relevant than those in the general statistical network: the RF outdegree positively influences the probability of success in four out of six models, the exception being full professorship in Methodological Statistics and Economic Statistics. In these models, in fact, being cited by colleagues of the same RF (RF_indegree), instead of having cited them (RF_outdegree) seems to influence the success probability, consistently with what is expected in a competition for a prominent position.

RF_indegree also improves the probability of success for associate professorship in Methodological Statistics, while general indegree has a negative effect; this seems to indicate a qualitative control on the disciplinary consistency of the research products, beyond the quantitative computation of its impact.

Table 9 Final results of logistic regressions for each statistics position and RF. Citation-based indices were included in the first variables block without any selection. ***, **, *, and °denote significance at the 0.1%, 1%, 5%, and 10% levels

	Parameter estimates					
	13/D1 AP	13/D1 FP	13/D2 AP	13/D2 FP	13/D3 AP	13/D3 FP
(Intercept)	− 1.947**	− 2.726**	− 1.901**	− 1.378	− 3.909**	− 4.811*
ND_I1	0.037	0.013	0.024	0.021	− 0.033	0.048
ND_I2	0.093	0.231***	0.265*	− 0.002	0.167	− 0.020
ND_I3	0.611	0.008	0.047	− 0.142	0.652	− 0.436
CD_I1	− 0.067	− 0.095°	− 0.149	− 0.019	0.150	− 0.099
CD_I2	− 0.003	− 0.003	− 0.011	− 0.001	− 0.002	− 0.000
CD_I3	0.234	0.300°	− 0.110	− 0.033	0.083	− 0.004
Accomplishments					0.381	0.623
Bonpow		0.545*				
Indegree	− 0.501**					
Self citations	− 0.042*				− 0.143*	
RF_indegree	0.728***		0.368*	0.396**	0.581***	
RF_outdegree	0.149**		0.409**			0.287*
Mc Fadden Pseudo R^2	0.283	0.208	0.272	0.189	0.392	0.306
Accuracy	0.7132	0.7593	0.717	0.7258	0.7778	0.8605

Self-citations negatively affect the probability of success for associate professorship both in Methodological Statistics and in Demography and Social Statistics, while Bonacich's power centrality measure turns out to be significant and positively associated with success only for full professorship in Methodological Statistics.

Citation network measures are quite relevant to explain the ASN success, especially RF indegree and RF outdegree.

Quantitative bibliometric indices When we add to the models quantitative bibliometric indices analogous to those used in the CDs, not much changes, and the models' accuracies only show a slight improvement, ranging from 71.7 to 86.0% (see Table 9). All the citation network variables that were significant in the previous models remain significant, with the only exception of the model for full professorship in Methodological Statistics, where the RF_indegree is no more significant after taking into account the h-index and number of published journal articles (both significant al 10% level). Apart from this model, no other model shows any significance of the newly added variables. We can just observe some minor differences: e.g., in some models the number or journal articles published in a list of top-class journals becomes significant once we control for the citation-based indices (it is the case with associate professorship in Methodological Statistics and Economic Statistics). Other minor changes relate to the significance of both RF_indegree and RF_outdegree for associate professorship in Economic Statistics (only the RF_outdegree was first significant), and the significance of the number of positively evaluated titles for associate professorship in Demography and Social Statistics.

Models including quantitative bibliometric indices show no relevant differences.

Discussion

In the ASN procedure, the appointed committee is requested to assess the candidates' CVs based on quantitative and qualitative characteristics. While the quantitative points of interest are well described in the Law 240/2010 (Law, 2011) and discussed in the literature related to ASN (see, e.g. Marzolla, 2016; Marini, 2017; Demetrescu et al., 2020; Abramo & D'Angelo, 2015), the main focus of the qualitative evaluation is more ambiguous and undefined. Nevertheless, qualitative analysis of CVs does not seem to have a minor role in the assessment; Marzolla (2015) reports that, in the first ASN session, only 52.8% of the over-median candidates obtained the qualification,[7] while Abramo and D'Angelo (2015) reveal a scarce capacity of the bibliometric performance alone to explain the success in the ASN procedure. Our analyses confirm the weak predictive power of the quantitative component of the evaluation: in models where success is predicted by bibliometric indices and the number of achieved accomplishments, accuracy only reaches about 60% for statisticians and 70% for computer scientists. The particularly low accuracy obtained with models fitted on the data on statisticians might depend on an "unofficial" adoption of citation-based metrics in the NDs; however, the introduction of citation-based indices in the models did not significantly improve the models.

This partial discrepancy between bibliometric performance and ASN attainment might appear strange and unexpected, but one should remember that bibliometric indices are the same whatever RF the candidate chooses to apply for. This implies that one of the basic tasks for the committee members, which cannot be achieved automatically, might be to evaluate if the candidate's scientific production is consistent with the scientific domain of the RF.

Citation networks are sometimes used to identify communities (Ji & Jin, 2016), i.e. groups of authors connected to each other by common topics and references, and could then be used to measure the degree of affinity of a candidate with an academic group. The academic group can be envisaged as the discipline in a broad sense, or as the RF. For this reason, citation networks in our application are somewhat different from a proper complete citation network, since they only involve citations among candidates and academics in a restricted subsample of the Italian university system, according to the broader or narrower definition.

Citation practices may vary markedly among disciplines, with an average number of citations per article ranging from less than one to more than six according to the discipline, but Computer Science, Mathematics and Social Sciences show similar values (Adler et al., 2009; Amin & Mabe, 2000). However, the two analysed citation networks, for computer scientists and for statisticians, are quite different in size, with both a number of candidates and a number of academics in Computer Science more than twice as large as in statistics. Statistics and Computer Science networks also differ in the citation network structure, with a more dense network for computer scientists compared to statisticians; this is partly due to a higher productivity of computer scientists, but might also indicate a larger tendency to cite Italian authors. The Computer Science citation network is also more centralised, while the statisticians' network is far more hierarchical, suggesting that the computer scientists' network shows a more clear distinction between very central authors and very peripheral ones, while statisticians are characterised by a strongly asymmetric network.

---

[7] In the first ASN session, medians were used as thresholds for each quantitative index.

Despite these structural differences, both disciplines share a strong correlation between network measures and ASN success [RQ1]: successful candidates show higher values of indegree, outdegree, self-citations and prestige, which means they are more cited by their colleagues, and they cite their colleagues more (and also themselves). This is not a mere indirect effect of a larger productivity and a better bibliometric performance. In fact, network measures significantly improve the model fit and affect the success probability, net of the ASN indices of scientific productivity [RQ2]. This is especially true with network measures based on the narrower RF network, in each model and in both disciplines, suggesting that, independently on the bibliometric performance, citing and being cited by colleagues who belong to the same research field leads to better results than citing or being cited outside the scientific clique. This is in line with the findings of Abramo and D'Angelo (2015), who report a larger probability of success for candidates already on staff compared to candidates external to academia or, among the incumbents, a better result for those classified in the same RF compared to those from other RFs. However, Abramo & D'Angelo interpret this result in terms of favouritism and discrimination, without considering the possibility of a consistency check of the candidate's research interests with the RF contents.

Things are similar, but not the same, for Computer Science and Statistics. A number of network measures contribute to explain the probability of success in both cases but, as regards statisticians, in five out of six models none of the ASN indices has a significant effect on the success probability. This does not imply that the committees do not check that the candidates exceed at least two out of three thresholds, but the quantitative measures of productivity do not seem to be as central in the final judgement as expected. Consequently, network measures are more relevant for the success probability of statisticians than of computer scientists, and in many cases are the only significant variables in the fitted models for the Statistics discipline [RQ3]. This holds true when controlling for the citation-based indices that are used in the Computer Science models: again the bibliometric performance variables are non-significant in almost all the models, and the network measures are the only predictive variables [RQ4].

The scarce relevance of the bibliometric performance variables to explain the ASN success for statisticians is likely to depend on the lower degree of maturity this discipline has reached in terms of bibliometric tradition; in fact, computer scientists show a higher awareness of citations and citation-based indices, as witnessed by the fact they publish more, they co-author more frequently, they cite each other more extensively.

Another reason for the larger impact of the network measures among statisticians compared to computer scientists has probably to do with the characteristics of the research fields pertaining to the disciplines. The perceived differences between 01/B1 and 09/H1 seem to be smaller than the difference among 13/D1, 13/D2 and 13/D3. The rate of applicants in the RF 09/H1 who also apply for the RF 01/B1 is much higher than the rate of candidates who apply to more than one of the statistical RFs (76.6% vs. 17.5%), signalling a larger overlap in the Computer Science discipline. We can derive similar considerations from the observation of the citation network measures calculated for the broader and narrower networks: when we move from discipline networks to RF networks, densities become larger, and the improvement is much bigger for the statistical RF. In other words, while computer scientists cite each other without respect to RF differences, it is not true that statisticians do not cite each other, but instead they cite those who belong to the same research field. Moreover, the RF networks for statisticians are more centralised than the general discipline network, suggesting that, when we focus on a more consistent group of

academics, it becomes as easy to discern central and peripheral authors as it is for computer scientists.

Conclusions and future work

In this paper, we investigated the role of citation networks in explaining the attainment of the national qualification in citation-based and non-citation-based research fields, namely in the Computer Science and in the Statistics disciplines. Citation network measures turned out to be relevant to predict the ASN qualification, particularly for the non-citation-based RFs in Statistics. Besides, the most predictive measures are referred to the RFs sub-networks: only citations to and from scholars from the same RF guarantee that the research interests are within the disciplinary boundaries of the research field. These boundaries seem to be more compelling for statisticians, in agreement with the lower degree of overlapping in the RFs applications. The additional information conveyed by citation network measures testifies to the recognition and/or influence exerted by scholars in the same discipline, and can be seen as an indicator of qualitative consistency, i.e., affinity with the specific scientific community.

Conversely, citation network measures are not a substitute for citation-based bibliometric indices in the Statistics RFs, since they do not lose their predictive power when the citation-based indices are forced in the models: in spite of the ongoing transition from non-citation-based to citation-based assessment, there seems not to be any implicit use of citation-based indices on the part of statisticians. One might conclude that this implies a still unripe and incomplete transition, but a longitudinal analysis of the network characteristics, and a comparison of the effects of each block of variables on the probability of success across time will be suitable to determine at what stage the transition is, and where it is heading. This might imply to move from the "candidate-perspective" that we adopted in this research to an "application-perspective", addressing the correlation issues raised by multiple applications to the same research field and role.

The large differences observed between the two disciplines of Computer Science and Statistics, in spite of their similarities in terms of citation practices, suggest that in the future it will be interesting to widen the analysis to more, and more different, research fields, such as sheer Humanities fields. Since books and book chapters are the main types of publication in the Humanities, this will force to address a series of problems of information retrieval and citation normalisation.

This analysis is based on citation networks, but Ji and Jin (2016) pointed out the complementarity of citation and co-authorship networks in identifying communities; another possible development of this research might be to enlarge the analysis to co-authorship networks.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

References

Abramo, G., & D'Angelo, C. A. (2015). An assessment of the first "scientific habilitation" for university appointments in Italy. *Economia Politica, 32*(3), 329–357. https://doi.org/10.1007/s40888-015-0016-9

Adler, R., Ewing, J., & Taylor, P. (2009). Citation statistics. *Statistical Science, 24*(1), 1–14. https://doi.org/10.1214/09-STS285

Amin, M., & Mabe, M. (2000). Impact factor: Use and abuse. *Perspectives in Publishing, 1*, 1–6.

Bedogni, L., Cabri, G., Martoglia, R., & Poggi, F. (2022). Does the venue of scientific conferences leverage their impact? A large scale study on Computer Science conferences. *Library Hi Tech*. https://doi.org/10.1108/LHT-09-2021-0305

Bendel, R. B., & Afifi, A. A. (1977). Comparison of stopping rules in forward "stepwise" regression. *Journal of the American Statistical Association, 72*(357), 46–53. https://doi.org/10.1080/01621459.1977.10479905

Bonacich, P. (1972). Factoring and weighting approaches to status scores and clique identification. *The Journal of Mathematical Sociology, 2*(1), 113–120. https://doi.org/10.1080/0022250X.1972.9989806

Bonacich, P. (1987). Power and centrality: A family of measures. *American Journal of Sociology, 92*(5), 1170–1182. https://doi.org/10.1086/228631

Borgatti, S. P. (2005). Centrality and network flow. *Social Networks, 27*(1), 55–71. https://doi.org/10.1016/j.socnet.2004.11.008

Butts, C. T. (2008). Social network analysis with sna. *Journal of Statistical Software, 24*(6), 06. https://doi.org/10.18637/jss.v024.i06

Dawson, S., Gašević, D., Siemens, G., & Joksimovic, S. (2014). *Current state and future trends: A citation network analysis of the learning analytics field*. 231–240. https://doi.org/10.1145/2567574.2567585

Demetrescu, C., Finocchi, I., Ribichini, A., & Schaerf, M. (2020). On bibliometrics in academic promotions: A case study in Computer Science and engineering in Italy. *Scientometrics, 124*(3), 2207–2228. https://doi.org/10.1007/s11192-020-03548-9

Di Iorio, A., Poggi, F., & Peroni, S. (2019). Open data to evaluate academic researchers: An experiment with the Italian Scientific Habilitation. In *17th International Conference on Scientometrics and Informetrics, ISSI 2019—Proceedings Volume 2*, (pp. 2133–2144). http://arxiv.org/abs/1902.03287

Diallo, S. Y., Lynch, C. J., Gore, R., & Padilla, J. J. (2016). Identifying key papers within a journal via network centrality measures. *Scientometrics, 107*(3), 1005–1020. https://doi.org/10.1007/s11192-016-1891-8

Ding, Y. (2011). Scientific collaboration and endorsement: Network analysis of coauthorship and citation networks. *Journal of Informetrics, 5*(1), 187–203. https://doi.org/10.1016/j.joi.2010.10.008

Ding, Y., Yan, E., Frazho, A., & Caverlee, J. (2009). PageRank for ranking authors in co-citation networks. *Journal of the American Society for Information Science and Technology, 60*(11), 2229–2243. https://doi.org/10.1002/asi.21171

Feeley, T. H., LaVail, K. H., & Barnett, G. A. (2011). Predicting faculty job centrality in communication. *Scientometrics, 87*(2), 303–314. https://doi.org/10.1007/s11192-010-0324-3

Freeman, L. C. (1978). Centrality in social networks conceptual clarification. *Social Networks, 1*(3), 215–239. https://doi.org/10.1016/0378-8733(78)90021-7

Goldberg, S. R., Anthony, H., & Evans, T. S. (2015). Modelling citation networks. *Scientometrics, 105*(3), 1577–1604. https://doi.org/10.1007/s11192-015-1737-9

Hastie, T., Friedman, J., & Tibshirani, R. (2001). *The elements of statistical learning*. Springer. https://doi.org/10.1007/978-0-387-21606-5

Jensen, P., Rouquier, J.-B., & Croissant, Y. (2008). Testing bibliometric indicators by their prediction of scientists promotions. *Scientometrics, 78*(3), 467–479. https://doi.org/10.1007/s11192-007-2014-3

Ji, P., & Jin, J. (2016). Coauthorship and citation networks for statisticians. *The Annals of Applied Statistics, 10*(4), 1779–1812. https://doi.org/10.1214/15-AOAS896

Krackhardt, D. (1994). Graph theoretical dimensions of informal organizations. In *Computational organization theory* (pp. 89–111). Lawrence Erlbaum Associates, Inc.

Law. (2011). Rules concerning the organization of the universities, academic employees and recruitment procedures, empowering the government to foster the quality and efficiency of the university system (Norme in materia di organizzazione delle università, di personale accademico e reclutamento, nonche' delega al Governo per incentivare la qualità e l'efficienza del sistema universitario). *Gazzetta Ufficiale n. 10 Del 14 Gennaio 2011 Suppl. Ordinario n. 11.* http://www.gazzettaufficiale.it/eli/id/2011/01/14/011G0009/sg

Liu, X., Jiang, S., Chen, H., Larson, C. A., & Roco, M. C. (2015). Modeling knowledge diffusion in scientific innovation networks: An institutional comparison between China and US with illustration for nanotechnology. *Scientometrics, 105*(3), 1953–1984. https://doi.org/10.1007/s11192-015-1761-9

Marini, G. (2017). New promotion patterns in Italian universities: Less seniority and more productivity? Data from ASN. *Higher Education, 73*(2), 189–205. https://doi.org/10.1007/s10734-016-0008-x

Martini, M. C., Pelle, E., Poggi, F., & Sciandra, A. (2021). *Datasets and results of the paper titled "Are citation networks relevant to explain academic promotions? An empirical analysis of the Italian national scientific qualification."* https://doi.org/10.5281/zenodo.6364742

Marzolla, M. (2015). Quantitative analysis of the Italian National Scientific Qualification. *Journal of Informetrics, 9*(2), 285–316. https://doi.org/10.1016/j.joi.2015.02.006

Marzolla, M. (2016). Assessing evaluation procedures for individual researchers: The case of the Italian National Scientific Qualification. *Journal of Informetrics, 10*(2), 408–438. https://doi.org/10.1016/j.joi.2016.01.009

Nuzzolese, A. G., Ciancarini, P., Gangemi, A., Peroni, S., Poggi, F., & Presutti, V. (2019). Do altmetrics work for assessing research quality? *Scientometrics, 118*(2), 539–562. https://doi.org/10.1007/s11192-018-2988-z

Pautasso, M. (2015). The Italian University Habilitation and the challenge of increasing the representation of women in academia. *Challenges, 6*(1), 26–41. https://doi.org/10.3390/challe6010026

Peroni, S., Ciancarini, P., Gangemi, A., Nuzzolese, A. G., Poggi, F., & Presutti, V. (2020). The practice of self-citations: A longitudinal study. *Scientometrics, 123*(1), 253–282. https://doi.org/10.1007/s11192-020-03397-6

Poggi, F., Ciancarini, P., Gangemi, A., Nuzzolese, A. G., Peroni, S., & Presutti, V. (2019). Predicting the results of evaluation procedures of academics. *PeerJ Computer Science, 5*, e199. https://doi.org/10.7717/peerj-cs.199

Radicchi, F., Fortunato, S., & Vespignani, A. (2012). Citation networks. In A. Scharnhorst, K. Börner, & P. van den Besselaar (Eds.), *Models of science dynamics: Encounters between complexity theory and information sciences* (pp. 233–257). London: Springer. https://doi.org/10.1007/978-3-642-23068-4_7

Scarpa, F., Bianco, V., & Tagliafico, L. A. (2018). The impact of the national assessment exercises on self-citation rate and publication venue: An empirical investigation on the engineering academic sector in Italy. *Scientometrics, 117*(2), 997–1022. https://doi.org/10.1007/s11192-018-2913-5

Singh, A. P., Shubhankar, K., & Pudi, V. (2011). *An efficient algorithm for ranking research papers based on citation network* (pp. 88–95). https://doi.org/10.1109/DMO.2011.5976510

Stephenson, K., & Zelen, M. (1989). Rethinking centrality: Methods and examples. *Social Networks, 11*(1), 1–37. https://doi.org/10.1016/0378-8733(89)90016-6

Tregellas, J. R., Smucny, J., Rojas, D. C., & Legget, K. T. (2018). Predicting academic career outcomes by predoctoral publication record. *PeerJ, 6*, e5707. https://doi.org/10.7717/peerj.5707

Vieira, E. S., Cabral, J. A. S., & Gomes, J. A. N. F. (2014). How good is a model based on bibliometric indicators in predicting the final decisions made by peers? *Journal of Informetrics, 8*(2), 390–405. https://doi.org/10.1016/j.joi.2014.01.012

Wallace, M. L., Larivière, V., & Gingras, Y. (2012). A small world of citations? The influence of collaboration networks on citation practices. *PLoS ONE, 7*(3), e33339. https://doi.org/10.1371/journal.pone.0033339

Wasserman, S., & Faust, K. (1994). *Social network analysis: Methods and applications* (1st ed.). Cambridge University Press. https://doi.org/10.1017/CBO9780511815478

Yan, E., & Ding, Y. (2009). Applying centrality measures to impact analysis: A coauthorship network analysis. *Journal of the American Society for Information Science and Technology, 60*(10), 2107–2118. https://doi.org/10.1002/asi.21128

Authors and Affiliations

Maria Cristiana Martini[1] · **Elvira Pelle[1]** · **Francesco Poggi[2]** · **Andrea Sciandra[1]**

Maria Cristiana Martini
mariacristiana.martini@unimore.it

Elvira Pelle
elvira.pelle@unimore.it

Andrea Sciandra
andrea.sciandra@unimore.it

[1] Department of Communication and Economics, University of Modena and Reggio Emilia, Modena, Italy

[2] Italian National Research Council (CNR), Institute of Cognitive Sciences and Technologies (ISTC), Rome, Italy