Studies in Higher Education

ISSN: 0307-5079 (Print) 1470-174X (Online) Journal homepage: https://www.tandfonline.com/loi/cshe20

Research performance evaluation: some critical thoughts on standard bibliometric indicators

Loukas N. Anninos

To cite this article: Loukas N. Anninos (2014) Research performance evaluation: some critical thoughts on standard bibliometric indicators, Studies in Higher Education, 39:9, 1542-1561, DOI: 10.1080/03075079.2013.801429

To link to this article: https://doi.org/10.1080/03075079.2013.801429

Studies in Higher Education, 2014
Vol. 39, No. 9, 1542–1561, http://dx.doi.org/10.1080/03075079.2013.801429

Research performance evaluation: some critical thoughts on standard bibliometric indicators

Loukas N. Anninos*

Department of Business Administration, University of Piraeus, 80 Karaoli and Dimitriou Str, 185 34 Piraeus, Greece

The bibliometric methodology is an established technique for research evaluation as it offers an objective determination and comparison of research performance. This paper aims to critically assess some standard bibliometric indicators commonly used (based on publication and citation counts) to evaluate academic units, and examine whether there are factors not taken into account that influence evaluation results. Findings suggest that the dissimilarity of academic units (for example regarding their scientific orientation and size in terms of staff number), if not taken into consideration may lead to incorrect evaluation results. This issue becomes even more important when comparing larger academic units, such as universities or colleges. Among the suggested further research is the replication of this study in an expanded time frame and an international context.

Keywords: university; ranking; research evaluation; performance; higher education

Introduction and objectives

University performance evaluation includes (a) systematic, documented and detailed performance measurement according to predefined dimensions/criteria and the institutional type/character, goals and mission; (b) decision making for improvements (wherever necessary) in the educational and organizational dimension; and (c) comparison (national or international in scope) of performance with similar academic units. Performance evaluation in universities may be at institutional, departmental or study programme level and includes qualitative and quantitative indicators of inputs, processes, outputs and outcomes (e.g. Cave et al. 1997; Hämäläinen et al. 2002; Katharaki and Katharakis 2010; Baldrige Education Criteria 2011–2012). According to Diamond (2002), performance evaluation allows the assessment of the credibility of goals, the suitability of academic strategy and policy, the introduction of changes and innovations, and the efficient and effective academic and organizational operation of institutions. In specific cases, performance evaluation may focus on a single academic process, such as teaching or research.

University research is the systematic process of investigating and studying within various scientific fields in order to reach new conclusions through the use of scientific methods. Quantification of research outputs through performance evaluation is an issue

*Email: lanninos@unipi.gr

of high significance for many industrialized countries, as university funding may depend on the quality of research performance (Bazeley 2010; Edgar and Geare 2013).

A methodology that has been developed for evaluating research performance is bibliometric analysis. The bibliometric analysis is based on the use of several quantitative parameters (thus it is objective under certain preconditions) to quantify the research results of an academic unit, and it is an objective and the most common research evaluation methodology (Toncich 2006). The indicators or systems used constitute the basis for ranking the evaluated academic units.

However, there is considerable scepticism expressed regarding the methodology, dimensions/criteria and weighting schemes used in the ranking approach (Provan and Abercromby 2000; Dill and Soo 2004; Van Dyke 2005; Marginson 2006, 2007; Buela-Casal et al. 2007; Salmi and Saroyan 2007; Usher and Savino 2007; Williams and Van Dyke 2007; Cheng and Liu 2008; Högskoleverket 2009; Rauhvargers 2011).

From the analysis of ranking systems, it is concluded that they constitute a superficial and inadequate assessment methodology which is far from marking out genuinely excellent universities. Each ranking system has a different view of university excellence and serves a different purpose (Van der Wende 2008). Their frequent methodological changes prevent comparative evaluations in the long term and do not allow improvements to be recorded. In addition, ranking systems do not offer comprehensive university evaluation (e.g. in the managerial and educational dimension). Universities with a strong focus on sciences that publish in English are usually ranked higher. However, the fundamental weakness of the ranking approach is that it often overlooks national priorities and strategies, differences in the mission and profile of institutions, and disciplinary specializations of academic units. So, the comparability of academic units remains an issue.

Bibliometric evaluation of research

Bibliometrics has been used globally for over 30 years to evaluate research performance of individuals, teams and academic units in a variety of scientific fields, to recognize interdisciplinarity and scientific collaboration, to delve into the structure of scientific fields, to reveal the relation between science and technology (e.g. publications and patents) and to document the impact of research on scientific progress (Narin 1976; Moed 2005). It is based on two hypotheses: (a) academics aim towards the progress of science, which means diffusion of scientific results through different communication channels; and (b) academics need to publish in order to build their reputation and evolve professionally (Archambault and Gangé 2004; Rehn and Kronman 2006).

Figure 1 shows that bibliometric evaluation of research begins with selection of the evaluation framework, the unit of analysis, time period and database or set of journals. Then, the type of data and data collection methodology along with limitations are determined. The evaluation process continues with the extraction of data from bibliographic databases (such as Thomson Scientific) and the calculation of simple and/or compound indicators of publications, citation and impact, scientific collaboration and journal quality. Academic units may then be ranked accordingly. The last step of this process involves the extraction of conclusions and relevant decision making.

However, this method has its own limitations. Figure 2 shows some limitations of bibliometric analysis (Borgman and Siegfried 1992; Archambault and Gangé 2004; Moed 2005; Van Raan 2005; Rehn and Kronman 2006; Campbell and Bertrand 2009).

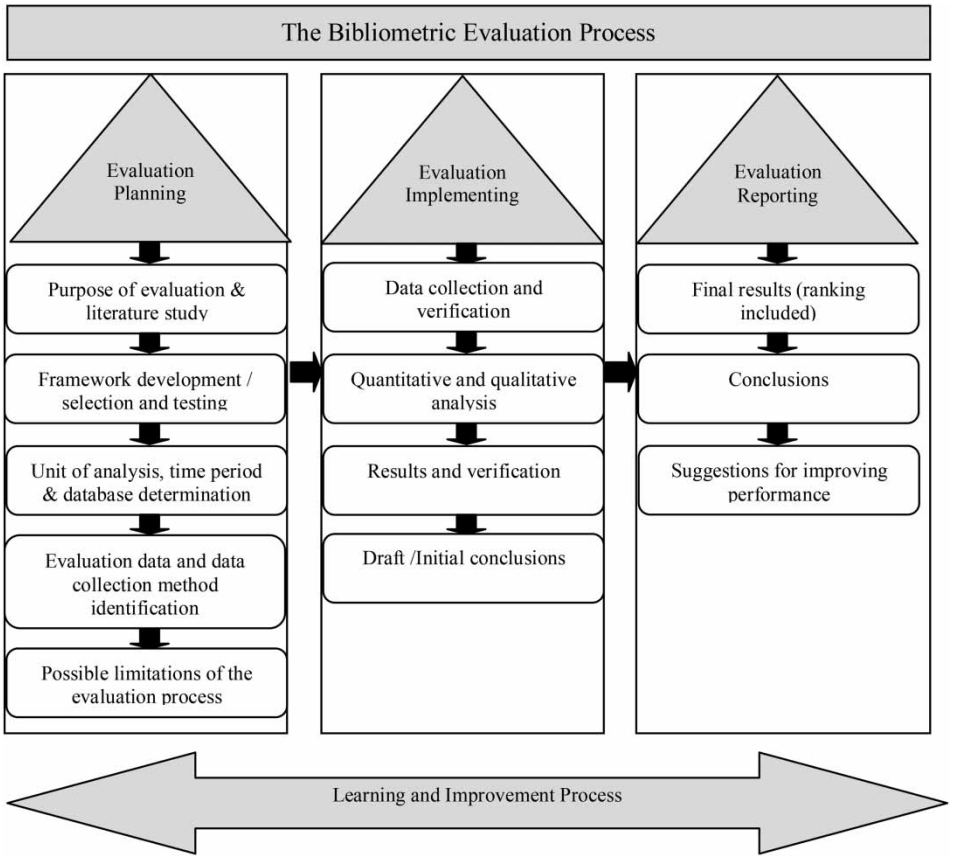

Figure 1. The process of evaluating research performance through bibliometrics.

- Limited journal coverage in bibliographic databases
- Journal classification to scientific fields
- Greater coverage of sciences in comparison to social sciences
- Changes in journal title
- Misspellings in the name, address and affiliation of authors and citations
- Inaccurate classification of publications to authors
- Non coverage of certain publication types (eg books)
- Excessive, negative citations, self citations and personal motives
- Languages coverage (bias for English)
- Differences in publications and citations patterns between fields

Figure 2. The limitations of bibliometric analysis.

This dynamic trend of using bibliometric data to evaluate research presupposes that institutions will make an effort (perhaps by developing evaluation-specific processes) to present research performance to interested parties.

Bibliometric evaluation of research is used for ranking universities, research groups and departments at the national and international level, in the sciences and social sciences sectors. The aforementioned limitations as well as the peculiarities of social sciences (differences in publication patterns and citation frequency, different width of research results, readership, journal coverage by relevant databases) influence evaluation results and should be taken into account by researchers that conduct evaluation studies. Table 1 presents some studies that attempted to evaluate research performance in various disciplines with the use of standard bibliometric indicators. The studies shown offer rankings based on simple or adjusted bibliometric indicators. Some of them examine the changes in ranking positions of academic units, and others present the consistency of results among similar studies in the same discipline. Some of the studies mention additional determinants of research performance (such as funding or size of academic unit). The methodologies used include the selection of a number of scientific journals, a time period, data collection and relevant analysis (publications, pages, citations, citations per paper, etc). In certain cases indicators can be weighted.

Despite the proliferation of ranking methodologies, their popularity and the ongoing attempts made by the ranking actors to produce a comprehensive ranking system, the departmental and institutional diversity (regarding mission, objectives, strategies, priorities, especially at an international scale) poses a serious query regarding the possibility of comparison and the extraction of accurate information on quality. In most cases, academic units are ranked according to their performance on each indicator without examining how similar and comparable they are to each other, thus producing ambiguous results. It is well known that universities which have medical schools or various science departments will have higher publication and citation counts compared to universities with schools of arts or departments of humanities. The same stands for academic units in which research is produced by scholars in science fields in comparison with units with a great number of social science scholars. Moed and Visser (2007) and Lopez-Illescas, Moya-Anegon and Moed (2011) suggest that university ranking systems should account for differences in their disciplinary specialization, something that is investigated in this paper on a departmental level.

This paper focuses on the problem of comparative evaluation and ranking of academic units. More specifically, this study offers a detailed bibliometric evaluation and ranking of two types of academic department in Greece, namely economics and management. While bibliometric evaluation and ranking of departments of economics and business have been reported in the past, this study approaches ranking results critically searching for factors not taken into account (such as disciplinary orientation and size in terms of staff number) that influence evaluation results and determine the position of academic units in ranking tables. The issues raised are of particular importance for bibliometrics and university ranking in general.

Methodology

Before any attempt to use bibliometric analysis to evaluate research performance, the suitability of this method for application to a specified setting should first be examined.

Table 1. Some studies that evaluate research performance using bibliometrics.

Author(s)	Year	Ranking	Finding(s)
Niemi	1975	Southern USA (economics 1970–1974)	Performance evaluation and ranking
Gerrity and McKenzie	1978	Southern USA (economics 1976–1977)	Use of citations as a measure of individual/departmental contribution
Graves, Marchand and Thompson	1982	US universities (economics 1974–1978)	Pages, pages per academic member Salaries and secretary/faculty member related to publications Teaching and service loads as inhibiting factors Size should be taken into account
Davis and Papanek	1984	US universities (agricultural economics and economics 1978, 1981)	Citation based ranking should account for age and dispersion
Laband	1985	US universities (economics 1971–1983)	Past performance is a proxy of future performance, citations are measures of quality, length of time employed permits the evaluation of member improvement, quality and quantity of published research by graduates and job placement are indicators
Blair et al	1986	US departments of economics 1977–1981	Citation per capita and citation per capita excluding self citations
Harris	1990	Australian departments of economics 1984–1988	Performance measurement according to publications, publication per staff member, proportion of staff cited, number of works cited, citations/staff member, per cited member Variation determinants: no. of lecturers, teaching load, student/staff ratio, secretary/assistant/computers per staff, internal/external funds
Nederhof and Noyons	1992	Dutch experimental psychology 1980–1985	The significance of the similarity of departments in terms of structure is presented
Nederhof and Van Raan	1993	Six research groups (economics 1980–1988)	Bibliometric analysis and peer review enriches the process of research evaluation
Kirman and Dahl	1994	European countries and universities (1987–1991)	Articles, authors per country, balance teaching and research differs among countries, bias in favour of English
Conroy and Dusansky	1995	US (economics 1987–1997)	Ranking based on number of pages, pages per faculty member
Lucas	1995	Canadian departments of economics 1981–1990	Total and per capita pages, adjusted for co-authorship and journal quality

(*Continued*.)

Table 1. (*Continued*.)

Author(s)	Year	Ranking	Finding(s)
Scott and Mittias	1996	US universities (economics 1984–1993)	Ranking based on number of pages, pages per faculty member, concentration; bias against smaller departments
Dusansky and Vernon	1998	US universities (economics 1990–1994)	Ranking based on number of pages, pages per faculty member
Noyons, Luwel and Moed	1998	Flanders (IT 1983–1992)	Performance evaluation of IT subdomains
Kalaitzidakis, Mamuneas and Stengos	1999	European institutions and countries (economics 1991–1996)	Comparison of performance highlighted
Jin and Yau	1999	East Asian universities (economics 1990–1996)	Pages, different journal sets do not dramatically change ranking, school history and tradition are not directly related to research productivity
Coupe	2000	European and US economists (1969–2000) and institutions (1990–2000)	The size of the academic unit should be taken into consideration when comparing performance
Rovira, Senra and Jou	2000	Catalonian universities (physics 1991–1998)	Papers, citations, CPP, subfields should be taken into account
Erkut	2002	Canadian universities (business schools 1990–1999)	Research output may be affected by scientific fields, culture, economic reasons
Kalaitzidakis, Mamuneas and Stengos	2003	World and European Universities (economics 1995–1999)	Ranking of economics departments based on adjusted pages and total pages
Lubrano et al	2003	European universities (economics 1991–2000)	Set of journal affects ranking, educational policy, promotion and funding may affect results
Visser, Nederhof, and Van Raan	2003	Erasmus University (management 1999–2001)	Performance evaluation of Erasmus University and 38 benchmarks in management and its subfields
Van Leeuwen et al.	2003	Netherlands (chemistry 1991–2000)	It is noted that small differences between universities are not taken into account in ranking
European Commission	2004	European universities Economics subfields (1991–2000)	Performance evaluation of countries and ranking of institutions
Moed	2006	European and US universities (1997–2004)	It is suggested that university profiles and disciplinary specialization should be taken into account so that appropriate benchmarks are chosen for comparative evaluation studies
Moed and Visser	2007	Netherlands (informatics 1996–2001)	It is suggested that any scientific scholarly research evaluation

(*Continued*.)

Table 1. (*Continued*.)

Author(s)	Year	Ranking	Finding(s)
			system has to take into account differences among subfields and scholarship
Jin and Hong	2008	East Asian universities (economics 1990–2005)	Ranking based on pages Suggested that rankings related only to economics profession, difficult to generalize to total university ranking
Noyons	2008	Lund University (2002–2007)	Performance evaluation of departments and research groups (and world comparisons)
Nederhof	2008	Dutch individuals and institutes of economics (1994–1998, 2000)	When the focus of a research performance system is on outputs only, it may increase the number of publications in the more highly weighed international outlets, but may fail to obtain a high citation impact
Markusova et al	2009	Russia (sciences 1997–2007)	Evaluation based on publications and citations
Courtault et al	2010	France Economics and management	Ranking suggested based on Hirsch Index
Bodkin	2010	Canadian Economics 1984–1996	Evaluation based on papers and citations

As suggested in the literature, the scientific fields of economics and business can be evaluated with the use of bibliometric indicators (however, not comprehensively, as not all means of scientific communication are included in bibliographic databases). It is important to know the percentage of national publications that are included in the databases that will constitute the basis for evaluation, as Greek academic units are evaluated. Since there is a significant share of research output included in the Thomson Scientific databases, bibliometric analysis can be applied.

The methodology of the present study is consisted of nine consecutive stages, as given in Table 2.

Table 2. Methodology stages.

(1)	Determination of academic level for evaluation and analysis
(2)	Determination of time period for analysis
(3)	Determination of databases for data extraction
(4)	Determination of publication types for analysis
(5)	Identification of possible limitations
(6)	Identification of departmental publications and counting method
(7)	Conduct of citation analysis
(8)	Calculation of bibliometric indicators
(9)	Presentation of results and critical analysis

In the first stage (*Determination of academic level for evaluation and analysis*) the level of analysis was determined. For this study, all Greek university departments of economics and business were selected for evaluation and analysis. This study differs from previous studies in Greece in terms of methodology, indicators and evaluation of results. The basic objective here is to examine how comparable are departments of the same discipline. This element is important when ranking departments and crucial when ranking universities.

Secondly, the *time period* that would set the limits for the bibliometric evaluation was set to 10 years (1996–2005), which according to the literature is sufficient for bibliometric analysis (Archambault and Gangé 2004; Moed 2005, 2006). However, as the primary scope of this study is the evaluation of specific indicators, the selection of this decade is only an intermediate step which serves the basic research objective.

Then, the *databases for data extraction and analysis* were selected. These were the Science Citation Index and Social Sciences Citation Index of the Web of Science produced by Thomson Scientific, which allow more comprehensive analyses (Hicks 1999; Archambault and Gangé 2004). Its major strong points include multidimensionality (all scientific fields are included), journal selectivity (based on impact), full journal coverage, author affiliation and address information, citations data and online availability of its content. This study is not restricted to a set of journals.

Then, *possible limitations* were identified, e.g. (a) the accuracy of data (author, affiliation information, errors in citation information), which largely depends on the Thomson Scientific databases; (b) the type of scientific output (articles, reviews and relevant citations); (c) only counting publications in journals covered by Thomson Scientific; and (d) the 10-year period for analysis for publications and citations. Also, this study does not take into account the personal characteristics and demographics of scholars or the national parameters that can influence research productivity and quality.

The next stage is *identification of publication types* that will be taken into consideration for the bibliometric analysis. Therefore, only publications that are denoted as articles or reviews are included in the analysis. These two publication types constitute the main types of scientific communication through journals. Other publication types such as editorial notes, abstracts, corrections, letters, comments and book reviews are not included.

The sixth stage includes the *identification of departmental publications* that fall into the selected time period (1996–2005) and the *counting method*. The stage begins with the identification of academic staff (professors, associate professors, assistant professors and lecturers) employed in each academic unit according to the official archives of the Greek Ministry of Education. Detailed lists indicating author names, affiliation and disciplinary specialization were provided. The publications that are included in the analysis (with one author or more) must have at least one author with a Greek name and include a Greek institution in the address field. Then, publications are counted based on the 'article method', which means that the number of publications for an academic unit depends on the number of times the unit is referred in the address field (whole counting). If there is no information about the affiliation of an author, the above process is performed. If no information can be found, the publication is not counted for analysis. It is noted however, that in cases where academics were employed at a different institution in the past (Greek or foreign), their publications are counted in their current affiliation.

The next stage in the methodology is a *citation analysis* of the above publications. Despite the fact that every publication in a refereed journal meets certain quality

Table 3. Bibliometric indicators.

Indicator Symbol	Description
P	Number of publications (articles, reviews) in journals included in THOMSON ISI (Science Citation Index, Social Science Citation Index)
C	Number of citations to publications in journals included in THOMSON-ISI (self citations are excluded).
C_{+sc}	Number of citations to publications in journals included in THOMSON-ISI (self citations are included).
CPP	Average number of citations per publication (self citations are excluded)
Pnc	Percentage of publications without citations during the predefined time period (self citations are excluded)
JCSm	Average citation number of all papers published in the journals in which an academic unit is active (journal set) (self citations are excluded). The JCSm is a reference value for constructing complex indicators.
FCSm	Average citation number of all papers published in the scientific fields in which an academic unit is active (self citations are excluded). The average citation number of each scientific field is considered as world average. The FCSm is a reference value for constructing complex indicators.
CPP/JCSm	Impact of the academic unit's papers compared to the average citation number of the unit's journal set.
CPP/FCSm	Impact of the academic unit's papers compared to the world citation average in the scientific fields in which the unit is active.
JCSm/ FCSm	Impact of the journals in which an academic unit has published, compared to the word citation average in the scientific fields covered by these journals.
% Self Citations	Percentage of self citations (citations in which the citing and the cited paper have at least one author in common).

criteria, no information is provided regarding the impact of the publication on the international scientific community. This impact is measured by the number of citations. A publication with a great number of citations is usually considered a quality publication. So, this stage aims at measuring all citations that refer to the publications of Greek economics and business departments. Citations are counted in the same way as publications (whole counting). It is also noted that only citations that are properly entered in the Thomson Scientific databases are taken into account.

The eighth stage includes the *calculation of the selected indicators* that have been used extensively in the past, as in Table 3.

After the calculation of the indicators has been completed, the economics and business departments whose research performance was evaluated are ranked in the final table. The final ranking results are then critically approached and analysed.

Results and discussion

The performance of each department is presented according to the bibliometric indicators used (Table 4). Departments of economics are indicated with the letter O and departments of management are indicated with the letter D.

Table 4. Evaluation results.

Department	P	C	C +SC	CPP	PNC	% SC	CPP/ JCSm	CPP/ FCSm	JCSm/ FCSm
O-1	120	243	301	2,03	69	19	0,83	0,8	1,04
O-2	103	79	117	0,77	65	33	0,53	0,32	0,6
O-3	98	77	106	0,79	61	27	0,41	0,25	0,61
O-4	71	55	70	0,77	66	21	0,4	0,22	0,54
O-5	47	49	87	1,06	57	44	0,51	0,45	0,89
O-6	32	26	35	0,81	66	26	0,49	0,37	0,75
O-7	32	28	37	0,88	59	24	0,47	0,34	0,73
O-8	25	43	55	1,72	56	22	0,81	0,54	0,67
O-9	13	11	11	0,85	69	0	0,99	0,49	0,5
O-10	8	14	14	1,75	25	0	0,93	0,71	0,76
O-11	79	126	153	1,59	49	18	0,58	0,54	0,93
O-12	38	35	44	0,92	74	20	0,79	0,51	0,64
O-13	11	22	24	2	18	8	0,9	1,06	1,18
O-14	2	0	0	0	100	0	0	0	0,48
O-15	69	78	108	1,13	58	28	0,53	0,36	0,67
O-16	21	22	30	1,05	43	27	0,58	0,36	0,63
O-17	13	34	37	2,62	46	8	0,98	0,74	0,75
D-1	80	155	214	1,94	46	28	0,67	0,56	0,83
D-2	38	103	123	2,71	53	16	0,75	0,7	0,94
D-3	31	69	76	2,22	39	9	1,35	0,84	0,62
D-4	17	6	8	0,76	71	32	0,65	0,38	0,59
D-5	14	44	52	3,14	47	15	0,64	0,47	0,73
D-6	10	33	37	3,3	10	11	0,84	0,76	0,91
D-7	9	3	9	0,33	89	67	0,28	0,22	0,77
D-8	132	257	334	1,95	54	23	0,76	0,71	0,93
D-9	46	89	104	1,93	57	15	0,9	0,67	0,74
D-10	11	5	9	0,45	64	44	0,46	0,31	0,68
D-11	40	72	103	1,8	48	30	0,75	0,6	0,8
D-12	19	16	21	0,84	58	24	0,66	0,37	0,57
D-13	15	13	19	0,4	67	32	0,16	0,11	0,74

It is simple to use bibliometric indicators to compare two (theoretically similar) departments of economics and reach a conclusion on which one is performing better. For example the department O1 has more publications than the department O6, therefore O1 is better than O6. A similar conclusion is drawn if two departments of business administration are compared. The department D8 is performing better in terms of citations in comparison to the department D5. However, these results are sustained only if O1 and O6 departments and D8 and D5 departments are indeed similar and comparable to each other. Apart from the specific characteristics of every academic unit (such as size in terms of academic staff), differences may be present in terms of research orientation (social sciences or sciences), type of publications, dominant publication patterns (papers in high or low impact journals, conferences, books, etc.), scientific

collaboration, etc. Faculty productivity can be influenced by personal characteristics (e.g. gender) and abilities, personality traits, motivation, background, age and disciplinary specialization (and relevant correlates such as journal acceptance rates and rate of knowledge obsolescence), while a number of institutional factors can be related to high or low productivity rates, e.g. the teaching and service load, available resources and rewards (Creswell 1985; Creamer 1998; Toutkoushian et al. 2003). These factors should be very carefully examined and taken into account in any attempt to compare academic units, as ranking indicators can be influenced to a greater or lesser extent.

Once compatible bibliometric indicators have been produced, the similarity of the evaluated academic units must be taken into account. Two of the factors that are expected to influence the research performance of academic units are the scientific orientation and staff number of departments. In this study, the scientific orientation of departments is examined on the basis of the main disciplinary specialization of academic staff (as indicated in national databases). Two main categories were created: sciences and social sciences. For example, a specialized scholar in mathematics/informatics is included in the science category, while a specialized scholar in sociology or education is included in the social sciences category. These differences among fields can influence the performance of departments even in the case of economics and business administration.

Figure 3 presents the scientific orientation of the evaluated academic departments. The figure shows that departments that theoretically belong to the same scientific field (e.g. business) have different scientific orientations. The synthesis of academic staff and their productivity influence research performance and the homogeneity of academic units. If most papers from a business administration department are published by academics whose research orientation does not fall into social sciences (informatics,

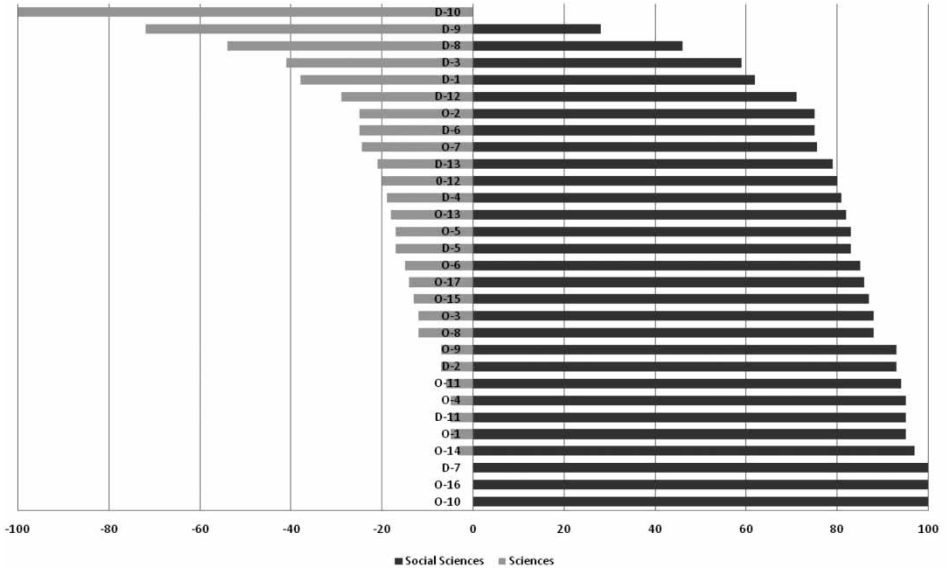

Figure 3. Scientific orientation of departments.

environment, etc.), this is something that impacts both the journals and the scientific fields in which the unit is active.

In addition, the variation in publication and citation frequency between different scientific fields should be considered. Different disciplines have different publication norms and citation patterns. Despite the fact that there are complex indicators which take into consideration the journals, fields, average citation rates based on time, and publication type, there is no information regarding the units' scientific orientations. From this angle a different view of the results is revealed. It is not correct to compare the citation per publication (CPP)/FCSm indicator between two departments (1.2 and 2 respectively), one of which has a 90% orientation to social sciences while the other has significantly lower orientation (30%). The comparison would be meaningful if both departments had comparable scientific orientations.

The same problem of comparability also exists when departmental size in terms of staff number is examined. A typical application of the publication indicator would suggest that a department with 15 academics and 40 publications is better than a department with 10 academic and 30 publications. If the number of staff is taken into account, the second department is better than the first one (under the hypothesis that they are comparable). Figure 4 shows publications per staff of all evaluated academic units.

Thus, application of the selected bibliometric indicators does not lead with certainty to sound conclusions regarding the potential superiority of one department over another. As it is uncommon for two academic units to be identical, in the case of a usual application of ranking systems, every unit should be individually evaluated. Thus, by examining each indicator separately, an initial attempt (mainly theoretical at this point) is made to identify the factors that affect them.

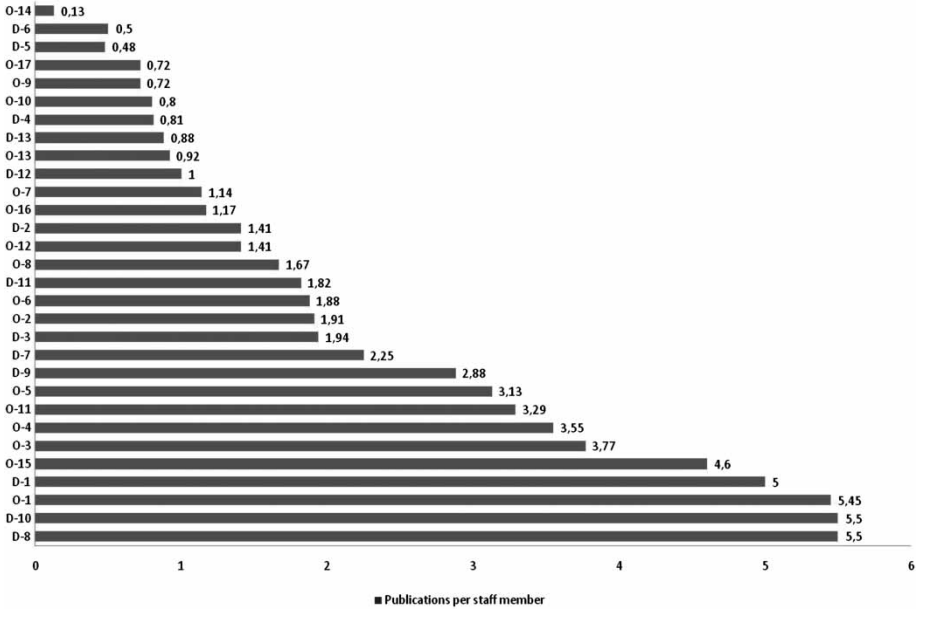

Figure 4. Publications per staff member.

Number of publications

The *number of publications* (P) in an academic unit may be influenced by the unit's size in terms of academic staff ($r = .488$, $df = 28$, $p < .01$). The first thing that should be examined is the size of an academic unit, as suggested by Toutkoushian et al. (2003) along with other researchers. Secondly, the percentage of academic staff whose research orientation falls into sciences or social sciences due to differences in publication and scientific communication patterns and frequency should be taken into account (in this case, the average publication number in departments oriented to sciences is 47.9 and that in departments oriented to social sciences is 40.1). Otherwise, science oriented departments are in favour. This can be explained by the fact that social science scholars have lower publication frequency mainly because they tend to publish in books (Pierce 1987; Clemens et al. 1995; Hicks 1999; Glänzel, Thijs and Schlemmer 2003; Archambault and Gangé 2004; Moed 2005). Hicks (1999) reports that 40–60% of literature in social sciences consists of books that attract 40% of total citations. But even within fields of the same discipline, there may be differences regarding their paradigmatic development, the understanding of accepted theory, the preferred methodologies and the understanding of important research areas (Creswell 1985). Publication rates are higher in 'high consensus' fields (mostly science) where the number of journals and acceptance rates are higher (Creamer 1998). Natural sciences are characterized by a higher acceptance rates among peer-reviewed journals, shorter length of papers and multiple authored papers (Lucas 1995).

In addition, the unit's mission and orientation to teaching or research should be examined as well as the type of scientific output covered by international databases. The number of publications at the individual level, apart from the subject and evaluation time frame is influenced by age and social status (Glänzel, Thijs and Schlemmer 2003), even if it decreases as the level of aggregation and heterogeneity increases. The synthesis of academic staff in professorial ranks should also be examined,

Number of citations

The second indicators, *number of citations excluding self-citations* (C) and *number of citations including self-citations* (C_{+SC}) are influenced by the publications' age, type, scientific field (which determines the frequency they are aggregated), and language. When these indicators are used for comparisons, the issue of publication age is not taken into account. Departments that have many publications in the past seem to achieve higher citation scores and are ranked higher. While this might be obvious, this is something that is not examined in the final ranking tables. In addition, as suggested earlier and according to the literature, citation frequency differs between disciplines. As a consequence, departments with many publications in scientific disciplines have higher citation rates.

Another important parameter that influences citation number is the type of publication. Publications that are reviews or experimental papers in sciences have higher frequency and higher numbers of citations because they are used as a basis for the evolution of science. Departments that have many of these kinds of papers seem to be performing better. In our study, there are only a few review papers. However, they attract a high percentage of citations. For example, department D5 has only two review papers that represent 52% of its total departmental citations. In department D2, the two review papers correspond to 26% of its total citations. Moreover, citation scores depend also on the year of publication. As a result, departments that have many old publications (reviews) seem to have higher citation scores.

The journals in which the research of academic units is published also play a role in the number of citations that papers attract. Papers in the top journals of any scientific field (as indicated through impact factors or other relevant indicators) have higher citation rates in comparison to papers published in journals of lower quality (for example, in the field of management, journals such as *MIS Quarterly* and the *Academy of Management Journal* have higher impact factors in comparison to *Total Quality Management and Business Excellence* and the *Canadian Journal of Administrative Science*. So, academic units that publish in top journals have higher citation scores. It is also to be remembered that authors are sometimes advised to include papers from the journal they wish to publish, thus forcing higher citations scores for journals.

Finally, the language of publication will determine if and how many citations will be gained (Van Raan et al. 2011). If a publication is written in English, it may be widely read and used. In the current study, no papers in Greek were identified, and therefore this factor has no influence over citation rates. In addition, cases of gratuitous authorship (something that influences the first indicator as well) and citations on papers that contain errors are also to be considered and examined.

Citations per publication

While the third selected indicator, *citations per publication* (CPP) is also influenced by publications' age, type, language and scientific field (for the same reasons as explained earlier), an interesting finding in this study was that the CPP values varied with the number of authors. This is something that has also been suggested by Huang, Xiaolan and Yuh-Shan (2008). More specifically, publications with more than three authors attract more citations than papers written by one or two authors. In total, 48% of publications are written by three or more authors and attain a CPP of 3.4, 34% are written by two authors and get a CPP of 1.58, and 18% are written by a single authors and attain a CPP of 1.7. This could possibly be explained by the multiplying value of knowledge. Scientific collaboration and interaction can offer insights from multiple angles, produce innovative insights and have an increased readership.

Publications without citations

The percentage of *publications without citations* (PNC) is influenced by the time period of the bibliometric analysis, the publications' scientific field, type and language, the specificity of knowledge, and whether the publications are considered by the international academic community to contribute to the evolution of science.

More specifically, the PNC indicator is influenced by the selected evaluation time frame. If the time frame is limited, the number of citations will be smaller. It is logical to suggest that when the time frame for evaluation is being decreased, the PNC value will be higher in the units oriented to social sciences. This happens because as mentioned earlier, the frequency of citation aggregation is lower in social sciences which means that departments that have many publications in social sciences tend to have higher PNC rates.

However, it should be noted that between different scientific fields there are publications that are very specialized (e.g. application of a less popular or novel technique to address a research problem). This means that these publications may have no or very few citations. It does not mean that are of inferior quality. Consequently, a high PNC rate is not necessarily a negative parameter because a large part of a unit's research output might be specialized.

In addition, there seems to be no relation between staff number and PNC ($r = .017$, $df = 28$, $p > .01$) based on the data of the present study. On the other hand, the PNC indicator marginally changes according to the age of publications.

The type of publications and language are theoretically expected to influence the PNC indicator. However, in the present study, the majority of publications (a) are of the article type; and (b) are written in English, and therefore their impact cannot be examined.

Self citations

The *self citation* (Sc) indicator in the publication of an academic unit should not be considered as a negative parameter. Scholars that conduct very specific research, work in new research areas and need to combine and refer to their research results are expected to have self citations in their publications. Nevertheless, such departments get a high Sc score and are ranked lower compared with other units. However, in order to examine whether high self citation scores occur due to innovativeness, research results, specialized research work or a practice adopted by many scientists, full text analysis should be applied. In this study, this indicator was not found to be correlated to staff number ($r = -.167$, $df = 28$, $p > .01$), publications ($r = .137$, $df = 28$, $p > .01$) or citations ($r = -.33$, $df = 28$, $p > .01$).

Compound indicators

The next three compound indicators (Citation per Publication / mean Journal Citation Score [CPP/JSCm], Citation per Publication / mean Field Citation Score [CPP/FCSm] and mean Journal Citation Score / mean Field Citation Score [JCSm/FCSm]) are based on citations (and hence they are influenced by the same factors that influence the number of citations). For CPP/JCSm there is a statistically significant correlation among the two constituents (CPP and JCSm) ($r = .787$, $df = 28$, $p < .01$), meaning that publication of papers in quality/high impact journals influences positively both the aggregation of citations and the CPP indicator. However, before we conclude that department A with CPP/JCSm = 0.8 is better than department B with CPP/JCSm = 0.6, the quality of journals should be examined in combination with the quality and type of publication. The third compound indicator, JCSm/FCSm, can provide this information, as it compares the average impact of journals to the average impact of the fields they belong.

Baring in mind the differences in the frequency of citations in sciences and social sciences, it appears in the present study that departments that are science oriented have higher CPP/JCSm scores than departments that are social sciences oriented. Having examined the normality of the CPP/JCSm indicator with the Kolmogorov–Smirnov test, hypotheses were formed to test the statistical significance of this difference. Then, the independent sample *t*-test was performed to identify differences between the top 10 departments oriented in sciences and the top 10 departments in social sciences. Results show that there is not a statistically significant difference between departments' means ($p = .619$). This finding, while contradictory, can be explained by the fact that departmental publications do not fall strictly in science journals and fields, something that could confirm a statistically significant difference between the two categories of departments. In addition, there is a statistically significant correlation between old publications (1996–2000), CPP ($r = .406$, $df = 28$, $p < .01$) and JCSm ($r = .541$, $df = 28$, $p < .05$). Based on current data, there is no statistically

significant correlation between old publications and CPP/JCSm. A possible explanation for that is that the age factor is levelled through the division of CPP with JCSm. Finally, the indicators CPP/JCSm and departmental staff number present no statistically significant correlation here ($r = -.349$, df = 28, $p > .059$).

In the case of the next compound indicator, *CPP/FCSm*, whose calculation is a time consuming process (as Thomson Scientific does not allow direct calculation of average citation rate in every scientific field) there are some similarities. There is a statistically significant correlation ($r = .69$, df = 28, $p < .01$) between the two constituents of the indicator (CPP and FCSm), meaning that research publications in scientific fields with high citation scores positively influence the CPP of the academic unit.

Again, having in mind the differences in the frequency of citations in sciences and social sciences, it appears in the present study that departments that are science oriented have higher CPP/FCSm scores than departments that are social sciences oriented. After examining the normality of the CPP/FCSm indicator with the Kolmogorov–Smirnov test, hypotheses were formed to test the statistical significance of this difference. Then, the independent sample *t*-test was performed for the top 10 departments oriented in sciences and the top 10 departments in social sciences. Results show that there is not a statistically significant difference between departments' means ($p = .758$). In addition, there is a statistically significant correlation between old publications (1996–2000), CPP ($r = .406$, df = 28, $p < .01$) and FCSm ($r = .58$, df = 28, $p < .01$). Based on current data, there is no statistically significant correlation between old publications and CPP/FCSm. Finally, the indicators CPP/JCSm and departmental staff number present no statistically significant correlation ($r = -.315$, df = 28, $p > .05$).

Finally, in the case of the third compound indicator (JCSm/FCSm), it can be suggested that the indicator will be higher for academic units that publish in journals with a higher frequency of citation in their field. There is a statistically significant correlation ($r = .89$, df = 28, $p < .01$) between the two constituents of the indicator (JCSm and FCSm), meaning that academic units that are active in scientific fields that have high citation average will also have high citation average in their respective journals. Therefore, the scientific orientation of departments in sciences and social sciences is expected to influence this indicator. In the present study it appears that departments that are science oriented have higher JCSm/FCSm scores than departments that are social sciences oriented. By following the same statistical procedure as earlier, results show that there is not a statistically significant difference between departments' means ($p = .957$).

Other important factors that are expected to impact this indicator are the type of publications and their age. As the majority of publications in this study are classified as articles, this factor cannot be examined here. By examining the existence of correlation between staff number and JCSm/FCSm, no statistically significant correlations were identified ($r = -.125$, df = 28, $p > .05$). The age of publication is correlated significantly to JCSm ($r = .541$, df = 28, $p < .01$) and FCSm ($r = .58$, df = 28, $p < .01$), which means that departments with many old publications have also higher journal and field citation scores in which they are active. Nevertheless, there is not a significant correlation among age and JCSm/FCSm ($r = .151$, df = 28, $p > .05$).

Conclusions

Based on the above findings, it is concluded that the attempt to describe and analyse ranking indicators and results should be made with cautiousness as there are factors that can influence indicators at various degrees. While these conclusions are useful

when comparing academic departments, in larger units of analysis, such as universities/colleges (especially at an international scale), these issues become even more complicated and the process of producing a reliable ranking is put into question. Hence, apart from the variety of definitions and views on the 'quality university' that should be highlighted, the issue of comparability is of equal significance for reaching sound conclusions on university performance evaluation.

Some further research issues include a further exploration of the combined effect of personal, disciplinary and institutional variables on research performance, the impact of the above-mentioned factors on each indicator, the replication of this study over an expanded time period and conducting a sensitivity analysis to examine the effect of a possible change of the aforementioned factors on each scientific field and final ranking results.

References

Archambault, E., and E.V. Gangé. 2004. *The use of bibliometrics in the social sciences and humanities*. Montreal, Quebec, Canada: Social Sciences and Humanities Research Council of Canada.

Baldridge Education Criteria. 2011–2012. http://www.nist.gov/baldrige/publications/upload/2011_2012_Education_Criteria.pdf.

Bazeley, P. 2010. Conceptualizing research performance. *Studies in Higher Education*, 35, no. 8: 889–903.

Blair, D.W., P.L. Cottle, and M.J Wallace. 1986. Faculty rating of major economics departments by citations: an extension. *The American Economic Review* 76, no. 1: 264–67.

Bodkin, R.G. 2010. Rankings of Canadian universities with regard to research in economics at the turn of the twenty-first century. *The Journal of Business Inquiry* 9, no. 1: 65–73.

Borgman, C.L., and S.L. Siegfried. 1992. Getty's synonym and its cousins: A survey of applications of personal name matching algorithms. *Journal of the American Society for Information Science* 43, no. 7: 459–76.

Buela-Casal, G., O. Gutierez-Martinez, Paz Bermundez-Sanchez, and O. Vadillo-Munoz. 2007. Comparative study of international academic rankings of universities. *Scientometrics* 71, no. 3: 349–65.

Campbell, D., and F. Bertrand. 2009. Bibliometrics as a performance measurement tool for the evaluation of research: The cases of the National Cancer Institute of Canada and the Canadian Forest Service. Paper presented at the 2009 CES Conference, Ottawa, 31 May–3 June, in Ottawa, Canada.

Cave, M., S. Hanney, M. Henkel, and M. Kogan. 1997. *The use of performance indicators in higher education*. London: Jessica Kingsley.

Cheng, Y., and N.C. Liu. 2008. Examining major rankings according to the Berlin principles. *Higher Education in Europe* 33, no. 2–3: 201–8.

Clemens, E.S., W.W. Powel, K. McIlwaine, and D. Okamoto. 1995. Careers in print: books, journals and scholarly reputations. *American Journal of Sociology* 101, no. 2: 315–30.

Conroy, M.E., and R. Dusansky. 1995. The productivity of economics departments in the US. Publications in core journals. *Journal of Economic Literature* 33, no. 4: 1966–71.

Coupe, T. 2000. *Revealed performances worldwide rankings of economists and economics departments*. Mimeo, ECARES Universite Libre de Bruxelles, Brussels, Belgium.

Courtault, J.M., N. Hayek, E. Rimbaux, and Z. Tong. 2010. Research economics and management in France: A bibliometric study using the h-index. *Journal of Socio Economics* 39, no. 2: 329–37.

Creamer, E.G. 1998. Assessing faculty publication productivity: Issues of equity. ASHE-ERIC Higher Education. Report 26 no. 2.

Creswell, J.W. 1985. Faculty research performance: Lessons from the sciences and the social sciences, ASHE-ERIC Higher Education. Report 4.

Davis, P., and G.F. Papanek. 1984. Faculty rating of major economics departments by citations. *The American Economic Review* 74, no. 1: 225–30.

Diamond, R.M., ed. 2002. *Field guide to academic leadership*. California: Jossey Bass.
Dill, D., and M. Soo. 2004. Is there a global definition of academic quality? A cross national analysis of university ranking systems, Public Policy for Academic Quality, Background Paper.
Dusansky, R, and C.J. Vernon. 1998. Ranking of US economic departments. *The Journal of Economic Perspectives* 12, no. 1: 157–75.
Edgar, F., and A. Geare. 2013. Factors influencing university research performance. *Studies in Higher Education* 38, no. 5: 774–92.
Erkut, E. 2002. Measuring Canadian business school research output and impact. *Canadian Journal of Administrative Sciences* 19, no. 2: 97–123.
European Commission. 2004. *Mapping of excellence in economics*. Brussels: European Commission.
Gerrity, D.M., and R.B. McKenzie. 1978. The ranking of southern economics departments. *Southern Economic Journal* 45, October: 608–14.
Glänzel, W., B. Thijs, and B. Schlemmer. 2003. A bibliometric approach to the role of author self citations in scientific communication. *Scientometrics* 59, no. 1: 63–77.
Graves, P.E., J. Marchand, and R. Thompson. 1982. Economics departmental ranking: Research incentives, constraints and efficiency. *American Economic Review* 72, no. 5: 1131–41.
Hämäläinen, K., K. Hämäläinen, A. Jessen, M. Kaartinen-Koutaniemi, and D. Kristoffersen. 2002. *Benchmarking in the improvement of higher education*. Helsinki: European Network for Quality Assurance in Higher Education.
Harris, G. 1990. Research output in Australian university economics departments: an update for 1984–1988. *Australian Economic Papers* 29, no. 55: 249–59.
Hicks, D. 1999. The difficulty of achieving full coverage of international social science literature and the bibliometric consequences. *Scientometrics* 44, no. 2: 193–215.
Högskoleverket. 2009. *Ranking of universities and higher education institutions for student information purposes?* Report Stockholm.
Huang, Y., A. Xiaolan, and H. Yuh-Shan. 2008. Use of citation per publication as an indicator to evaluate pentachlorophenol research. *Scientometrics* 75, no. 1: 67–80.
Jin, J., and J. Hong. 2008. East Asian rankings of economics departments. *Journal of Asian Economics* 19, no. 1: 74–82.
Jin, J., and L. Yau. 1999. Research productivity of the economics profession in East Asia. *Economic Inquiry* 37, no. 4: 706–10.
Kalaitzidakis, P., T.P. Mamuneas, and T. Stengos. 1999. European economics: An analysis based on publications in core journals. *European Economic Review* 43, no. 2: 1150–68.
Kalaitzidakis, P., T. Mamuneas, and T. Stengos. 2003. Rankings of academic journals and institutions in economics. *Journal of the European Economic Association* 1, no. 6: 1346–66.
Katharaki, M., and G. Katharakis. 2010. A comparative assessment of Greek universities efficiency using quantitative analysis. *International Journal of Educational Research* 49, no. 4–5: 115–28.
Kirman, A., and M. Dahl. 1994. Economic research in Europe. *European Economic Review* 38, no. 3–4: 505–22.
Laband, D. 1985. An evaluation of 50 ranked economics departments: by quantity and quality of faculty publications and graduate students placement and research success. *South Economic Journal* 52, no. 1: 216–40.
Lopez-Illescas, C., F. Moya-Anegon, and H.F. Moed. 2011. A ranking of universities should account for differences in their disciplinary specialization. *Scientometrics* 88: 563–74.
Lubrano M., S.C. Bauwens, A. Kirman, and C. Protopopescu. 2003. Ranking economics departments in Europe: a statistical approach. *Journal of the European Economic Association* 1, no. 6: 1367–1401.
Lucas, R. 1995. Contributions to economics journals by the Canadian economics profession, 1981–1990. *Canadian Journal of Economics* 28, no. 4: 949–60.
Marginson, S. 2006. Global university rankings: private and public goods. Paper presented at the 19th Annual CHER conference, September, 7–9, in Kassel, Germany.
Marginson, S. 2007. Global university rankings: where to from here? Paper presented at the Asia-Pacific Association for International Education. National University of Singapore, March 7–9, in Singapore.

Markusova, V.A., V.V. Ivanof, and A.E. Varshavskii. 2009. Bibliometric indicators of Russian science and of the Russian Academy of Sciences. *Herald of the Russian Academy of Sciences* 79, no. 3: 197–204.

Moed, H.F. 2005. *Citation analysis in research evaluation*. The Netherlands: Springer.

Moed, H.F. 2006. *Bibliometric rankings of world universities*. Center for Science and Technology Studies. Leiden University, the Netherlands. CWTS Report 2006–07.

Moed, H.F., and M.S. Visser. 2007. *Developing bibliometric indicators of research performance in computer science: An exploratory study*. Leiden, the Netherlands: Centre for Science and Technology Studies, Leiden University.

Narin, F. 1976. *Evaluative bibliometrics: The use of publication and citation analysis in the evaluation of scientific activity*, NJ: Cherry Hill CHI. In W. Glänzel. 2003. *Bibliometrics as a research field, Course Handouts*, pp. 6–11. http://www.cin.ufpe.br/~ajhol/futuro/references/01%23_Bibliometrics_Module_KUL_BIBLIOMETRICS%20AS%20A%20RESEARCH%20FIELD.pdf

Nederhof, A.J. 2008. Policy impact of bibliometric ranking of research performance of departments and individuals in Economics. *Scientometrics* 74, no. 1: 163–74.

Nederhof, A.J., and E.C.M. Noyons. 1992. Assessment of the international standing of university departments' research. A comparison of bibliometric methods. *Scientometrics* 24, no. 3: 393–404.

Nederhof, A.J., and A.F.J. Van Raan. 1993. A bibliometric analysis of six economics research groups: A comparison with peer review. *Research Policy* 22, no. 4: 353–68.

Niemi, A.W. 1975. Journal publication performance during 1970–1974: the relative output of southern economics departments. *South Economic Journal* 42, no. 1: 97–106.

Noyons, E.C.M. 2008. *Bibliometric study of Lund University*. Report. Center for Science and Technology Studies (CWTS), Leiden University, Leiden, The Netherlands.

Noyons, E.C.M., M. Luwel, and H.F. Moed. 1998. Assessment of Flemish R&D in the field of information technology. *Research Policy* 27, no. 3: 285–300.

Pierce, S. 1987. Characteristics of professional knowledge structures: Some theoretical implications of citation studies. *Library and Informational Science Research* 9: 143–71.

Provan, D., and K. Abercromby. 2000. *University league tables and rankings: a critical analysis*. Commonwealth Higher Education Management Services, Paper No. 30.

Rauhvargers, A. 2011. *Global university rankings and their impact*. European University Association, Brussels, Belgium.

Rehn, C., and U. Kronman. 2006. Bibliometric handbook for Karolinska Institutet, Karolinska Institutet. http://ki.se/content/1/c6/01/79/31/bibliometric_handbook_karolinska_institutet_v_1.05.pdf

Rovira, L., P. Senra, and D. Jou. 2000. Bibliometric analysis of physics in Catalonia: Towards quality or consolidation? *Scientometrics* 49, no. 2: 233–56.

Salmi, J., and A. Saroyan. 2007. League tables as policy instruments: Uses and misuses. *Higher Education Management and Policy* 19, no. 2: 31–68.

Scott, L.C., and P.M. Mittias. 1996. Trends in rankings of economics departments in the US: an update. *Economic Inquiry* 34, no. 2: 378–400.

Toncich, D.J. 2006. Key factors in postgraduate research. Brighton: Chrystobel Engineering, http://www.doctortee.net/files/PHDBK2006-00p.pdf.

Toutkoushian, R.K., S.R. Porter, C. Danielson, and P.R. Hollis. 2003. Using publications counts to measure an institution's research productivity. *Research in Higher Education* 44, no. 2: 121–48.

Usher, A, and M. Savino. 2007. A global survey of university ranking and league tables. *Higher Education in Europe* 32, no. 1: 5–15.

Van der Wende, M. 2008. Rankings and classifications in higher education: A European perspective. *Higher education: Handbook of theory and research*, 23: 43–71, J.S. Smart, ed. The Netherlands: Springer.

Van Dyke, N. 2005. Twenty years of university reports cards. *Higher Education in Europe* 30, no. 2: 103–24.

Van Leeuwen, T.N., M.S. Visser, H.F. Moed, T.J. Nederhof, and A.F.J. Van Raan. 2003. The holy grail of science policy: Exploring and combining bibliometric tools in search of scientific excellence. *Scientometrics* 57, no. 2: 257–80.

Van Raan, A.F.J. 2005. Fatal attraction: conceptual and methodological problems in the ranking of universities by bibliometric methods. *Scientometrics* 62, no. 2: 133–44.
Van Raan, A.F.J., T.N. Leeuwen, and M.S. Visser. 2011. Severe language effect in university ranking: particularly Germany and France are wronged in citation based rankings. *Scientometrics* 88, no. 2: 495–98.
Visser, M.S., A.J. Nederhof, and A.F.J. Van Raan. 2003. *Bibliometric profiles of management research at the Erasmus University 1999–2001*. Report ERIM, the Netherlands.
Williams, R., and N. Van Dyke. 2007. Measuring the international standing of universities with an application to Australian universities. *Higher Education* 53, no. 6: 819–41.