University of Connecticut
DigitalCommons@UConn

Economics Working Papers | Department of Economics

9-1-2007

Academic Rankings with RePEc

Christian Zimmermann
University of Connecticut

Follow this and additional works at: http://digitalcommons.uconn.edu/econ_wpapers

Recommended Citation
Zimmermann, Christian, "Academic Rankings with RePEc" (2007). *Economics Working Papers.* Paper 200736.
http://digitalcommons.uconn.edu/econ_wpapers/200736

This is brought to you for free and open access by the Department of Economics at DigitalCommons@UConn. It has been accepted for inclusion in Economics Working Papers by an authorized administrator of DigitalCommons@UConn. For more information, please contact digitalcommons@uconn.edu.

University of Connecticut

Department of Economics Working Paper Series

Academic Rankings with RePEc

Christian Zimmermann
University of Connecticut

Working Paper 2007-36

September 2007

341 Mansfield Road, Unit 1063
Storrs, CT 06269–1063
Phone: (860) 486–3022
Fax: (860) 486–4463
http://www.econ.uconn.edu/

This working paper is indexed on RePEc, http://repec.org/

Abstract

This documents describes the data collection and use for the computation of rankings within RePEc (Research Papers in Economics). This encompasses the determination of impact factors for journals and working paper series, as well as the ranking of authors, institutions and geographic regions. The various ranking methods are also compared, using a snapshot of the data.

Journal of Economic Literature Classification: A14, A10, A11, A13, Z00

Keywords: RePEc, rankings, impact factors, working papers, h-index, citations.

This paper benefited from discussions and electronic correspondence with Kit Baum, Oded Galor, Bill Goffe, N. Gregory Mankiw, and Ekkehard Schlicht. The data used in these rankings would not exist without the major contributions of Jose Manuel Barrueco Cruz, Kit Baum, Sune Karlsson, Thomas Krichel, Ivan Kurmanov and all the other volunteers working on RePEc

1 Introduction

RePEc has now become an important bibliographic service for Economics and related fields. A considerable amount of data has been collected regarding who authored what, where it was published, who reads it and where it is cited. One way to use this wealth of data is to compute rankings of individuals, journals (and series), institutions and even countries. Along with the growth of the underlying data, these rankings, even though they are still experimental, have grown in importance in the profession. Indeed, there is evidence that they are more and more used for evaluation purposes (promotion and tenure decisions). Also, country specific rankings have been used in various professional publications and even popular press.

It is therefore high time for the methodology behind these rankings to be explained. While a criterion like the number of citations may appear to be simple, it is necessary to understand how it is computed. Indeed, for ranking purposes in RePEc, self-citations are not counted, but citations to other versions of an articles are counted. It is also important to understand how the citations are found, i.e., what citations can be considered in the statistics.

Compared to other ranking exercices, the present one includes also some criteria that are unique, such as those based on readership and those based on the number of authors citing. It is also rare to find the same source being used both to establish impact factors of publications and rankings of authors or institutions. Finally, no other effort has included working papers, which have now become a very important way to disseminate research in Economics, if not the most important.

The RePEc project would never have been possible without the efforts of the many volunteers that have participated in one way or the other: the maintainers of the so-called RePEc archives who contribute the basic bibliographic data, all those who have contributed through their programming skills, through making available hardware and/or bandwidth, through giving advice or simply through spreading the word about RePEc. RePEc is committed to honor the work of these volunteers by making sure their work will never be subject to a fee, both for publishers and users, and will remain in the public domain.

The rest of the paper is structured in the following way. Section 2 describes how the various components of the data used in the rankings are gathered. Section 3 details the construction of the impact factors, Section 4 how articles and working papers can be ranked. The various criteria used to rank authors are introduced in Section 5. It discusses as well the various ways these criteria can be aggregated and justifies the choices made for the "official" rankings. Sections 6, 7 and 8 present the procedures to rank respectively institutions, geographic regions, and finally other rankings. Section 9 takes a snapshot of the data and documents the concordance of the the various rank criteria. Section 10 discusses how RePEc rankings differ from oter rankings. Section 11 concludes.

2 Data Gathering

The section describes how all the data is gathered to obtain the source for the rankings. All data comes from RePEc and other projects related to RePEc. As this data is continuously updated, the ranking are refreshed as well on a monthly basis.

2.1 Bibliographic Data

The source of all the bibliographic data is RePEc. RePEc (Research Papers in Economics, http://repec.org/) was founded in June 1997 under the leadership of Thomas Krichel as a followup project to NetEc, founded in 1993. Under very little central management, publishers (commercial or academic) contribute the bibliographic data (called metadata) themselves using a common format. This data is provided through their own servers, where anybody can retrieve it and use it. Thus RePEc is just a scheme to organize metadata and make it available to the public domain.

At the time of this writing, over 800 archives were contributing metadata to RePEc, thus covering: 1800+ series with 220'000 working papers, 640 journals with 280'000 articles, 1600 book chapters, 1000 books and 1500 software components, for a total of over 500'000 items. Almost 400'000 of them are available for download in full text.

So-called RePEc services are then allowed to use this data to display the data to public, provided this is done freely. Several websites display directly the data collected through RePEc, the most popular being IDEAS (http://ideas.repec.org/), EconPapers (http://econpapers.repec.org/), Inomics (http://inomics.com/), and finally Socionet (http://socionet.ru/).[1] An email notification service for new online working papers is also available (NEP, http://nep.repec.org/). Finally, data gathered by RePEc is relayed through the Open Archives Initiative and therefore made available even more widely, but to services that do not specialize in Economics, such as Google Scholar, Oyster, etc.

2.2 Author Data

For any ranking, one needs to collect information about the publications of an author. One great difficulty is the many ways an author's name may be indexed. For example, John Maynard Keynes may be listed in the bibliographic metadata as:

1. John Maynard Keynes
2. John M. Keynes
3. John Keynes
4. J. M. Keynes
5. J. Keynes

[1] NetEc, with its child projects WoPEc and BibEc, also used to display RePEc data. NetEc closed as it was not worth the maintenance effort given that competitors within RePEc were offering a superior product. Econlit also uses RePEc data for working papers through an exchange of services agreement with RePEc.

6. Keynes, John Maynard
7. Keynes, John M.
8. Keynes, John
9. Keynes, J. M.
10. Keynes, J.

and one an imagine many other ways, including misspellings. Variations are even more numerous if nicknames, titles or suffices (Jr., Sr., III) are used, or if accents are used. In addition, several people may have the same name, especially if the first name is abbreviated. Thus, an automatic attribution of works to the respective authors is bound to have high error quotes. Human intervention is necessary here.

The best people to perform this are the authors themselves. To do this, they register with the RePEc Author Service at http://authors.repec.org/. In doing so, they provide contact details, their affiliations (see next section), and their name variations expected in the metadata. The search engine then suggests them works from the RePEc metadata that match the name variations, works that the author then can add to their profile.

One may ask why authors would go through that trouble. There are several incentives (Krichel and Zimmermann 2005). First, without a registration, an author is not ranked and his research output does not count towards the ranking of the institutions he is affiliated with. Second, when registered, any authors obtains notification of new citations that are found within RePEc, a compilation of all citations as well as a detailed ranking analysis every month.

At the time of writing, over 14'000 authors were registered, claiming over 270'000 works as theirs, a little less than half the works listed in RePEc.

The RePEc Author Service is based at the University of Connecticut and is monitored by the author of this paper. It runs on open source software written by Ivan Kurmanov and financed by a grant from the Ford Foundation.

2.3 Institution Data

Institutional data is based on the institutional records collected in EDIRC (Economics Departments, Institutes, and Research Centers in the World, http://edirc.repec.org/). Since 1995, this website collects links to academics institutions and government agencies that would principally employ economists. The data is quite accurate, for example it lists within a university all relevant departments (economics, finance, agricultural economics, business schools, sometimes public policy and similar departments), research centers, institutes, formal research groups and some chairs, as long as economists form a substantial part or the staff, or economic issues are prominent in the mission of the group. A second condition is that this listed entity have its own website. It does not need to have its own server (virtual or not), but it needs to have a web page that is more substantial than just a listing of classes: there should be at least a listing of faculty by name.

Entities not based in universities can also be listed. The obvious one are central banks and government agencies directly applying economic policy, say ministries of

finance, treasury, labor, and industry, but also statistical agencies and various research agencies. The same applies to international organizations. Finally, independent research institutes and think tanks are also listed, but not most commercial institutions (banks, consultancies). The only exceptions are those that have a RePEc archive or that provide substantial research for free through their website. Associations and societies are also listed.

All in all, over 10'000 institutions are listed. If they are specialized in a particular field, they are categorized, and almost all governmental agencies are categorized. Institutions are also categorized by countries or, in the case of the United States, by state. When authors register with the RePEc Author Service, they have the opportunity to specify with which institutions they are affiliated with among those listed in EDIRC (except associations and societies), but they can also suggest new entities. If they do not fit with with the criteria of EDIRC, they are still kept in their list of affiliations.

EDIRC is housed and managed at the University of Connecticut by the author of this paper.

2.4 Citation Data

Citation counts are often considered to be the most useful metric of the impact of a piece of research. Finding citations is, however, not a trivial matter. It can be either performed manually at great cost, or automatically where it needs considerable fine tuning and many exception rules.

All citation data for rankings are provided by the CitEc project, http://citec.repec.org/, managed by José Manuel Barrueco Cruz, librarian at the University of Valencia. CitEc runs on hardware provided by the Technical University of Valencia.

CitEc downloads all papers in pdf format it can find, typically those that are not hidden behind a password or some IP protection. Those pdf files are then successively converted to PostScript and text. The text is then parsed to recognize the references, which are then paired with items listed in RePEc with a fuzzy matching algorithm on titles and authors. To prevent erroneous attributions, the level of confidence for a match needs to be set quite high. For somewhat lower levels of confidence, registered authors have the option to check and add appropriate citations.

At the time of this writing, over 120'000 documents have been processed, extracting over two million references, over one million of which refer to over 160'000 items listed in RePEc. Given that only freely available documents can be analyzed, a large part of those documents are working papers. This has advantages and disadvantages. Working papers are typically more recent than published articles, thus it allows to have a much more up-to-date analysis than with articles alone. However, citations in published articles are considered to be much more valuable than in working papers (erroneously, as discussed further in a subsequent section). This is somewhat corrected in two ways: 1) some commercial publishers provide directly information to CitEc about references in their articles; 2) for authors who have both the working paper and published article version of an item in their profile, the references found in a version can be attributed to the other.

2.5 Abstract Views and Downloads Data

Another measure of the impact of research is how often it has been "looked at". Abstract views statistics assess the attractiveness of the title, the authors or the general topic. Downloads statistics indicate how much the above and abstracts have contributed to attractiveness of the downloaded document.

Keeping track of abstract views is not difficult using the logs of a web server. The only drawback is that abstract displayed during uses of the search engine cannot be counted. Downloads are more difficult, given that they typically link to external servers. Thus some mechanism needs to be put in place to keep track of downloads.

The decentralized nature of RePEc complicates the compilation of these statistics. The participating services needs first to keep appropriate logs, and second they need to make them available in an appropriate format. The LogEc project, http://logec.repec.org/, managed by Sune Karlsson at Örebro University, tries collect this information. The following RePEc services provide information for downloads and abstract views: EconPapers, IDEAS, NEP and Socionet. The defunct NetEc also used to provide data. Other services that use RePEc data, in whole or part, unfortunately do not provide statistics. Among them are Inomics, Econlit, Oyster and any service making use of the RePEc data made available though the Open Archive Initiative (Google Scholar, for example).

Quite obviously, these statistics are subject to manipulation, as one could repeatedly download a paper to increase its count. For this reason, various information about the abstract viewer or downloader are recorded to prevent repeat counts. This is mainly performed through the use of the IP address, taking also into account IP clusters. Also, and this is mostly relevant for abstract views, visits by search engine robots need to be discarded as they not represent human readership. Some robots identify themselves, and they can easily be taken care of. Others do not obey standard protocol and need to be recognized as robots. Various identification mechanisms are used to filters these additional robots from the data. Complete details on how all this performed cannot be given here to preserve the accuracy of the data. But overall, about 75% of abstract views are thus discarded, less for downloads.

Whether is it an over-count or under-count of the true counts is unknown. Some robots may slip through. Some downloads are discarded as repeated despite originating from different users because they came from the same IP clusters. This happens in particular with institutions using a single cache server. We hope, however, that the numbers are sufficiently high for such accidents to even out relatively smoothly across all documents.

In addition, various checks and balances are implemented to recognize abnormal behavior, mostly from authors trying to manipulate the statistics. Obviously, it is not revealed how they are done, but let it be known that several authors have been caught.

Despite all these adjustments, LogEc records over two million abstract views and half a million downloads a month, or every document's abstract is viewed five to six time a month, and every item available online is downloaded once or twice a month, on average.

2.6 Further Refinements of the Data

As the works covered in RePEc contain both publications and pre-publications, there is an issue with several versions of the same work being listed. In particular, a working paper may appear in several series. Thus, for any measure that considers the numbers of works someone has authored, one should count distinct works. For technical reason, the matching of different versions is done only for works that are listed in a registered author's profile. The basis is a very similar title and the author's recognition of authorship.

Note that such works may have been cited in their different versions. A citation to any version is counted towards all versions. The same applies to references.

2.7 Discussion of Coverage

Quite obviously, only journals and working paper series that are listed in RePEc can be classified, and only authors that registered themselves can be treated. There are thus omissions. For example, the *Journal of Human Resources* or the *Brookings Papers on Economic Activity* are not considered at this time. This is obviously avoidable, but the structure of RePEc puts the burden of indexing on the publishers. Unlisted authors can easily correct this by registering themselves.

Being listed is not sufficient. The listing needs to be maintained, i.e., new items added as they are published. Some publishers are better at this than others, be it with regard to timeliness, completeness (missing items), coverage (years covered) or data quality (syntax errors, confusing author names). Again, it is up to the publishers to do their work. And registered authors also need to maintain their profile with any additions.

Deceased authors are kept in the database, but their affiliations are removed, the logic being that they cannot contribute to the academic life of their employer anymore. The RePEc Author Service maintenance team tries to keep their profiles current.

Note that while some journals present in other studies are not classified here, our rankings cover also working paper series that are typically neglected by other studies. There are also a limited number of chapters and books. It turns out that some working paper series have very high impact factors, while many journals have low impact factors. It is thus wrong to believe that research only is valued if it is published in a journal. More on this later, in the discussion of impact factors.

3 Computation of Impact Factors and Ranking of Series or Journals

Many ranking exercises for institutions or authors rely heavily on impact factors calculated elsewhere, and these impact factors are usually the most controversial issue with these rankings. Here we take a different approach in that the impact factors are determined with the RePEc data. We compute four sets of impact factors.

3.1 Simple Impact Factors

The computation of this simple impact factor is rather straightforward. Just find all citations to items in that particular series or journal, count those citations and divide by the number of items in the series or journal. Several adjustments are performed to the number of citations: 1) self-citations within the series or journal are discarded, to prevent self-inflation. Self-citations by authors are still counted, though. 2) Considering that a work may have appeared in different series, all versions of the cited and citing work are considered, but only one is counted. This matters as for example an article may be cited, while its working paper version is not, but the working paper series is still credited with this citation.

3.2 Recursive Impact Factors

Recursive impact factors are computed in the same way as the simple impact factors, except that every citation carries some weight. That weight is the recursive impact factor. It is thus the fix point of a function that could be specified in the following way:

$$R_I = \frac{1}{\sum_{\forall I} \sum_{i \in I} R_I} \frac{\sum_{c_J \in I} R_J}{\sum_{i \in I} 1} \quad \forall I$$

where R_I is the recursive impact factor of series or journal I, which has items i. c_J represents all citations from journal J. To guarantee that a fix point exists, the weights are normalized such that the average item (article or working papers) has a recursive impact factor of one. Also, when there are several versions of a citing items, the one with the highest impact factor is considered.

These factors are computed by iteration. In the first pass, simple impact factors are used, and then in each pass the recursive impact factors from the previous iteration are taken. This does, however, never converge completely, as new items and citations are continuously added to the database. The results are relatively stable, though. Concretely, the weights are recomputed everyday for all series and journals that are refreshed on IDEAS, that is those that have had any amendments in the bibliographic data and those that have not been refreshed for thirty days.

The recursive impact factor computed here is similar to the Google PageRank (Brin and Page 1998), that ranks web pages higher if they are linked to a lot, even more so if it is by web sites that have a high PageRank. The difference is that Google computes a different factor for every page, whereas we compute one for every journal or paper series. The idea of the PageRank is to determine the probability that a web surfer clicking randomly would end up at that page. In our case, this would be the probability, or rather something proportional to it, that a reader following randomly references in articles and papers would end up with a particular journal or working paper series.[2]

[2] Strictly speaking this would only be true if we did not account for different versions of the same item. Also, the reader would need to follow all citations, as the impact factor is not divided by the number of cited items. Some versions of PageRank do this, however.

3.3 Discounted Impact Factors

This factor is similar to the simple impact factor, with one important difference. Each citation counts for the inverse of the age in years (plus one) of the citing paper. Thus, if an article is cited in a paper dated in 2004 and we are in 2007, this citations would count for 0.25.

Such a factor gives an edge to what is cited now, and therefore highlights the publications series that are hot now. It does, however, not mean that its most recent publications are well cited, only that some of them, possibly old, are well cited now.

3.4 Recursive Discounted Impact Factors

This factor is the recursive version of the discounted impact factor. It thus uses its own factors as weights, multiplied by the age factor. This highlights publications series currently well cited in series that are currently well cited.

3.5 Discussion

Some other published impact factors differentiate by type of article, for example by giving different weights to full articles, notes and book reviews. One may also want to discard corrigenda. The metadata does not contain the type of the article and the title in the vast majority of the cases does not allow to infer the type. We thus abstract from these considerations.

Also, some journal issues are different. For example, the *American Economic Review* has one issue a year with non-refereed short articles, the Papers and Proceedings of the annual meeting of the American Economic Association. These short papers are less likely to be cited and add to the article count, thereby diluting the impact factor of the regular article. One could isolate these special issues, but the task then becomes subjective as other journals are subject to the same issues at varying degrees. We want to stay objective in our ranking and thus do not adjust. In this particular example, the American Economic Association does not want this distinction to be made anyway.

There are also some small sample issues. Some working paper series especially have few items and may as a results have unexpectedly high or low impact factors, high if just one item is often cited. The current solution is not to rank series or journals with fewer than 50 items. The impact factors are, however, used as is.

4 Ranking of Works

There are six different ways to rank works (working papers, articles, chapters, books). One is to simply count the number of citations it has gathered, again adjusting for different versions of the same item. The second is to discount each citation by its age. The remaining four are to weigh those citations by the impact factors of the citing series or journals.

Thus, if one were to add up all citations to articles in a particular journal, then divide the result by the number of articles, one would obtain the simple impact factor

(except that self-citations within the journal need to be excluded). Or if one were to add up the scores of all articles in a journal, with scores using the recursive impact factors and excluding self-citations, one would obtain the recursive impact factor. Doing this with simple impact factors would result in the factors of the first pass in the recursive impact factor computation.

5 Rankings of Authors

Every person registered in the RePEc Author Service with works listed in the profile is ranked. There are many ways to rank authors and this section discusses those used in the RePEc rankings. The strategy to aggregate the various rankings in then discussed.

5.1 Criteria Based on the Number of Works

The simplest of all ways to ranks authors is by the number of works they have authored. However, as working papers are also considered, the same work may appear several times, in different versions. These duplicates can therefore not be considered. A ranking including the duplicates is provided, but it is not used in the calculation of the aggregate rankings.

The number of distinct works thus serves as basis for the following criteria. They are a combination of simple counts and counts with weights from the simple or recursive impact factors with those counts divided by the number of authors or not. Thus, the following criteria are used (with their respective labels in bold face):

1. **NbWorks**: Simple count;
2. **DNbWorks**: Count divided by number of authors on each work;
3. **ScWorks**: Count with simple impact factor weights;
4. **AScWorks**: Count with simple impact factor weights divided by number of authors on each work;
5. **WScWorks**: Count with recursive impact factor weights;
6. **AWScWorks**: Count with recursive impact factor weights divided by number of authors on each work.

The two first criteria merely indicate how prolific an author is. The four others measure one characteristic of the quality of one's work: where it was published. It is an imperfect measure, given one may simple ride on the tails of other papers published in the same series or journal that have been frequently cited. But such count based solely on the impact factors are the ones most frequently used, as they do not necessitate the compilation of citations if one simply takes the impact factors from somewhere else.

Note that the discounted impact factors and recursive discounted impact factors are not used here. They could also be considered, but this would put too much weight on criteria based on the number of works in the overall rankings.

5.2 Criteria Based on Citation Counts

Here, we have criteria similar to those based on the count works, but we count citations. Self-citations are eliminated, and they may not be weighted, or weights by any of the four impact factors. And all these criteria may be divided by the number of authors or not.

In addition, we provide the h-index introduced by Hirsch (2005). His definition: *A scientist has index h if h of his/her N_p papers have at least h citations each, and the other $(N_p - h)$ papers have no more than h citations each*. Thus, this author would have at least h^2 citations (at least h papers with at least h citations each). Such a criterion puts more emphasis on an important body of work, instead of a few very highly cited papers, by giving higher score to those who have many cited papers. This index was developed for physics, where scientists write a lot of papers and also cite rather generously. Some physicists have h above 100, but in Economics it is very rare to have an h above 20, mainly due to the fact that economists write fewer, but more involved papers.

Finally, to criteria count the number of authors citing a particular author, first a simple count, second a considering the rank of the citing author, giving more points for highly ranked citers. This can measure how widely an author is cited. For example, this penalizes those that cite each other repeatedly ("citing clubs"). Note that each co-author counts for these criteria.

Thus, we have the following criteria based on citations:

1. **NbCites**: Simple citation count;
2. **ANbCites**: Citation count divided by number of authors on each work;
3. **ScCites**: Citation count with simple impact factor weights;
4. **AScCites**: Citation count with simple impact factor weights divided by number of authors on each work;
5. **WScCites**: Citation count with recursive impact factor weights;
6. **AWScCites**: Citation count with recursive impact factor weights divided by number of authors on each work;
7. **DCites**: Citation count discounted by age;
8. **ADCites**: Citation count discounted by age and divided by number of authors on each work;
9. **DScCites**: Citation count with discounted impact factor weights;
10. **ADScCites**: Citation count with discounted impact factor weights divided by number of authors on each work;
11. **WDScCites**: Citation count with recursive discounted impact factor weights;
12. **AWDScCites**: Citation count with recursive discounted factor weights divided by number of authors on each work;

13. **HIndex**: h-index;

14. **NCAuthors**: Count of citing registered authors;

15. **RCAuthors**: Rank weighted count of citing registered authors.

Due to scheduling difference between the upload of new citations and the ranking computations, which happens about a week later, the new citations are included for a minority of the authors in current ranking, but they are for all authors in the next issue of the rankings. Also, all self-citations by the author are of course excluded.

5.3 Criteria Based on Journal Page Counts

The following criteria only concern journal articles. Whether one publishes a note, which is shorter, or a full length article is an indication how editors feel about the contribution of an article. Also, some argue that editors allow particularly good pieces to run longer, while less important works are cut. Thus the page count can be an indication of the worth of one's publication record. Again, the page count can be weighted, or not and divided by the number of authors or not.

1. **NbPages**: Simple page count;

2. **ScPages**: Page count divided by number of authors on each work;

3. **WSCPages**: Page count with simple impact factor weights;

4. **ANbPages**: Page count with simple impact factor weights divided by number of authors on each work;

5. **AScPages**: Page count with recursive impact factor weights;

6. **AWScPages**: Page count with recursive impact factor weights divided by number of authors on each work;

Thus publishing a long article in an obscure journal is valued highly with the two first criteria, but barely factors in with the four others. Note that these are criteria that, contrarily to the others, pertain to a subset of all documents (articles). Also, this criteria can sometimes be somewhat misleading. For example, if a journal does not provide page numbers, either because they are missing in the metadata or because the article is online only and not in a paginated format, the number of pages defaults to one. This is justified by the fact that in some cases only the number of the starting page is provided, with is indistinguishable from a one page article. In addition, these criteria do not take into account the size of the pages. Some journals publish in A4 or Letter format, whereas most have smaller formats. Font size may vary as well, thus actual content of a page could be quite different from a journal to the other. No such adjustments are performed as there is no way to systematically verify those parameters and how they may change through the years, except through intensive manual labor that would count the average number of words per page, or something of that order.

Note also that the discounted impact factors are not considered. Adding them would be giving more weight to publications in journals. Given that many journals have impact factors lower than working paper series, there is no particular reason to privilege journals. Let the market decide what the better publication outlet is.

5.4 Criteria Based on Popularity on RePEc Services

Here, we measure how many times document abstracts have been viewed and how often they have been downloaded. As described in the section on LogEc, these statistics pertain to the subset of RePEc services that report such statistics. Furthermore, as all the metadata collected by RePEc is in the public domain, one cannot track how much it is used. But looking at the collected subset can still give good indications. Note that these statistics are checked for multiple views or downloads, and robot and web spider activity is excluded, as described above.

Again, we provide statistics with the criteria divided by the number of authors or not. Thus the following four criteria are available in the category:

1. **AbsViews**: Total abstract views in the past 12 months;
2. **AAbsViews**: Total abstract views per author in the past 12 months;
3. **Downloads**: Total downloads in the past 12 months;
4. **ADownloads**: Total downloads per author in the past 12 months.

Statistics are computed for the last 12 months. On the one hand, including a longer period allows the smooth out inherent short-term variability, for example new papers announced through NEP get a large one time boost, and authors may not yet have claimed them in their profile. On the other hand, the period considered should not be too long. First, this allows to take into account what is popular now, second it corrects for bias stemming from items having been listed for a long time, while even older material may have been added only recently.

Note that the basis for counting abstract views and downloads is the item (article, paper, etc.), and these numbers are aggregated for registered authors. Thus, when a author creates a profiles, the statistics for his/her papers are added also for the period where he/she was not registered.

For computational reasons, the criteria with statistics per author are computed with a one month delay.

5.5 Aggregation of Criteria

Quite obviously, with so many criteria, it is difficult to agree on who the best economists are, especially as the rankings certainly do not correlate perfectly. Some way to aggregate the rankings is required and unfortunately different ways of doing so give different results. In fact, they emphasize different aspects that all have some relevance. We discuss here some of them and then discuss our choice.

5.5.1 Harmonic Mean of Ranks

The harmonic mean is defined as

$$M_{-1} = N \frac{1}{\sum_{i=1}^{N} \frac{1}{r_i}},$$

where r_i is the ranking of an author in criterion i. In such a mean, very good rankings have a lot of weight, for example the first rank counts twice as much as the second one. But a one rank difference carries very little weight for higher numbers. This mean therefore rewards those who are particularly good in some category, and perhaps too much. For this reason, the harmonic mean is dampened somewhat by adding a constant to each rank, and then subtracting it from the mean.

5.5.2 Arithmetic Mean of Ranks

This is the easiest and most frequently used way to aggregate criteria and create indices. It is defined as

$$M_1 = \frac{1}{N} \sum_{i=1}^{N} r_i.$$

Doing poorly on one criterion penalizes an author particularly hard. Doing particularly well on one criterion to compensate is much more difficult. Thus, the arithmetic mean rewards those who rank consistently across criteria.

5.5.3 Geometric Mean of Ranks

The geometric mean is defined as

$$M_0 = \left(\prod_{i=1}^{N} r_i \right)^{\frac{1}{N}},$$

where $\prod$ symbolizes the product. The geometric mean penalizes poor rankings and emphasizes good rankings. To see this, notice that the geometric mean is the exponential of the arithmetic mean, and thus it dramatizes the features of the latter. Or put in another way, given a generalized mean with exponent p defined as

$$M_p = \left(\frac{1}{n} \sum x_i^p \right)^{\frac{1}{p}},$$

the geometric mean correspond to $p = 0$, which is between the arithmetic mean ($p = 1$) and the harmonic mean ($p = -1$).

5.5.4 Lexicographic Ordering of Ranks

This aggregation method is the one that rewards to the most extreme ranking well in a particular criterion. For an author, all ranks are ordered from best to worst, then all authors are ranked in the following way: first all those with their best rank being a first rank, the tie breaker being their second best rank, than third best. Once all authors with ranks one are exhausted, those with rank two as their best rank are taken, etc. This is akin to the ordering of words in the dictionary, hence it is named "lexicographic." This concept is also used in Economics to describe some preferences in utility theory.

5.5.5 Graphicolexic Ordering of Ranks

This method takes the lexicographic method, but turns it on its head, hence its newly-coined name: authors are ranked first by their best worst rank, then their best second worst rank to break ties, etc. This rewards authors that do not have a slip-up according to some criterion.

5.5.6 Sum of Percent of Best in Criterion

All the aggregation methods above only consider how someone is ranked according to the various criteria, but not far apart the rank ranks are for each criterion. For example, barely being first is valued in the same way as being first with a large advance on the second. One way to take the latter into account is to attribute 100% to the first ranked, and then proportionally percentages to the lower ranked authors. All these scores are then added. These priviledges those that have criteria where they are siginificantly better than others, especially for criteria where the dispersion of scores is larger.

5.5.7 Exclusion of Extremes

The truncated mean excludes the x largest and smallest values. This reduces the impact of outliers. In particular, if one thinks that the particular aggregation mean one has chosen is too much influenced by such outliers, using truncation can make the mean more credible. There is no particular guideline to choose what the value of x should be. An alternative is the Winsorized mean, where the truncated criteria are set to the rank of the largest respectively lowest ranks remaining.

5.5.8 Discussion and Aggregation Choice

We have identified 32 different criteria for ranking authors, and could have easily added more. In addition, we presented six aggregation methods, which can even be varied with the number of extremes to exclude and some other degrees of freedom. Each of the criteria can be multiplied by some weight. This is a dismaying array of possibilities, but we need to make choices. Those choices are easier if the criteria or aggregation methods lead to similar results. To some extend they do, as we see in a subsequent section, but there are noticeable differences. We still need to make a choice, take a stand.

Everyone would probably favor a combination of criteria and aggregation method that would favor oneself. We need to find something that is credible, in the sense that a person outside the profession would find it agreeable. We want to highlight the particular achievement, say that an author is particularly successful in downloads despite not having published much (yet), or that an author elicited many citations despite not being prolific. The harmonic mean achieves this, but needs to be tempered somewhat, and we thus add a constant of one to each rank. Also we include all criteria but one, the simple number of works (without distinguishing the distinct works), in the aggregation. For each author, we further truncate by dropping the best and worst ranking. Thus, in summary: we consider 29 of 31 possible rankings with aggregation through an adjusted harmonic mean.

These choices can, and should, be argued and we leave the reader the opportunity to try other ways to rank on the website[3].

6 Ranking of Institutions

When registering, each author has the opportunity to affiliate oneself to some institutions. For those that are listed in EDIRC, they the affiliation is recorded with an identifier which can be used to aggregate all authors from that institution. This allows subsequently also to rank institutions.

A few rules apply. Only institutions listed in EDIRC are ranked. An author can affiliate oneself to several institutions and all receive full credit for the author. If an institution is a sub-entity of another institution also listed in EDIRC, the latter receives also credit, and the former's score is computed, but it is not ranked (it does not increment the rank count). For each criterion, the institution's score is just the sum of the scores of each affiliated author. The only exception is the h-index, see below.

Quite obviously, institutions with many authors are advantaged. Clearly, taking an average score within an institution would make little sense, as author registration is not mandatory, and potentially lower ranked authors may be discouraged to register. On the contrary, adding up all authors' scores gives the right incentive: everyone should register, including students who already have authored something in RePEc.

One controversial aspect, though, is that for authors with multiple affiliations, each one counts equally and fully. Typically, an author is affiliated with a department, where one is employed, and may have a secondary affiliation with a research center (associations and societies are excluded). That secondary affiliation currently gets the same credit and the main one. For institutes with many secondary affiliates (like the NBER and CEPR, for example), this adds up quickly and puts them far out of reach. The obvious solution would be to let authors attribute a percentage to each affiliation, the system is unfortunately not designed to do so at the moment.

Finally, we need to explain how the h-index is computed in the case of institutions. Remember that for authors h is defined as the number of works with at least h citations. For institutions, we follow Schubert (2007) and define the institutional h as the number of authors affiliated to that institution with an h-index of at least h. As the h can only

[3] http://ideas.repec.org/cgi-bin/newrank.cgi

be an integer and the support of its distribution is even smaller as for authors, there are numerous ties. To break them, we adapt Ruane and Tol (2007). They augment h by a rational number between zero and one measuring the distance to the next h-index considering how many citations are required to reach it. In our case, we measure a similar distance, but considering how many authors with appropriate h-indices are necessary to reach the next step.

7 Ranking of Geographic Regions

To rank geographic regions (countries, US states), the same logic is used as for ranking institutions. All authors affiliated with institutions in a particular region are added to the pool of that region. However, authors with multiple affiliations have their scores split among all regions.[4] For example, if an authors is member of a department in California and institutes in California and Italy, two thirds on the score count towards California and one third towards Italy.

For authors with affiliations not listed in EDIRC, the geographic location of their affiliation is guessed from the address of its web page. If it still could not be found, then the home page of the author and then the email address are used. Obviously, this can still fail, as addresses with .com, .net, .org or .info are not geographically informative. But at least we tried[5].

Once all these attributions are made, we simply add up the scores, properly weighted. The only exception is again the h-index, where the same scheme as for institutions is used. Note that we do not calculate scores for the United States as a whole, as it would obviously be number one in every aspect. Rankings for every state are given, though.

8 Other Rankings

A wealth of data is available, and this allows us to establish various other rankings. A few examples are below, and more will be added once sufficient critical mass is present to display somewhat credible results.

8.1 Ranking in Geographic Regions

Once authors have been attributed to a particular region, it is easy to rank them within that region as well. The same applies to institutions within that region. Publishing rankings with very few entities or authors do not make much sense, though. For this reason, a minimum of five authors or five institutions need to be present. In some regions, there is little hope for authors to be listed, whatever their prestige, due to lack of participation by others in RePEc, or in small countries, due to the lack of economists. Therefore, rankings for regional conglomerates are presented as well, say the Mountain states in the US, Central America and the Caribbean, or Africa.

[4]The handling of multiple affiliations is different from institution rankings to avoid an author being counted several times in a region.

[5]Some errors are unavoidable. For example, at the time of this writing, the Pacific island nations of Niue and Nauru are ranked thanks to two authors using courtesy domains from these micro-nations.

At this point, rankings in geographic regions are just an excerpt of the worldwide ranking that aggregates all criteria, that is, there is no change in order from the worldwide to the regional classification. It is planned to change this in the following way: compute for each criterion the regional ranking, and then aggregate those regional rankings.

8.2 Ranking of Women Economists

Women are, unfortunately, quite underrepresented in the Economics profession. It appears, from a limited investigation, that they are further underrepresented within RePEc. One can still try to make a meaningful ranking with data collected within RePEc. Unfortunately, an author registering with RePEc does not declare his or her gender. This needs to be inferred from the first and middle names using a name data bank. There are, however, several difficulties: some names may be used for both females and males, and this may vary by culture. Also, given the international nature of RePEc, there is a incredible diversity in first names.

The following rules are applied for gender attribution: if there is more than 90% confidence the gender is correct, it is so attributed. The ambiguopus ones and the unrecognized once are then manually entered in exception tables, one for names that were not in the original tables, the other for case by case attributions[6]. In the end, only 0.35% are left without a gender. Just under 14% are identified as female.

The ranking of female economists is performed solely among female, that is, without considering the gender wide ranking: females are ranked according to each criterion and them the rankings are aggregated. This makes is possible that the order may be different from the classification of female among all economists.

8.3 Ranking in Fields

When registering, authors do not declare a field of research. It is therefore difficult to classify them within each field, although one could try to infer it from the JEL codes attached to their papers. However, as it is customary to put several JEL codes on each paper, and only about 21% of all papers have such a code, infered field attributions would not be reliable. However, an attribution of authors to fields is currently in development using data collected in the NEP project, http://nep.repec.org/.

For institutions, this is another story, as those that predominantly work in a particular field are coded as such within EDIRC. For example, institutions working in agricultural economics or finance are well identified. Also, certain institution types are well documented: central banks, think tanks, international organizations. For others, patterns in their names (or their English translation) are used. This is the case for Economics departments and business schools.

As for regional rankings, field rankings are just an excerpt from the worldwide ranking. It is planned to compute self-contained rankings in the future.

[6]Thanks to many authors for putting a picture of themselves on their web page!

9 A Glimpse at Results

We do not want here to give detailed rankings, there are available at http://ideas.repec.org/top/ and updated every month. In the following, we present a comparison of the various criteria and aggregation methods using a snapshot of the data on September 16, 2007, with 13946 authors registered affiliated to 3393 institutions.

9.1 Impact Factors

How do the impact factors compare? Table 1 provides a summary with rank correlations. All of them are very high. This is quite natural as series with many citations ought also to be cited by series with high impact factors. Overall, it does not seem to matter which criterion is used when it comes to rankings series or journals.

Looking only at the top 100 series (Table 2) correlations are reduced: the disparities between the top and worst series do not count anymore. This is reinforced when one does not filter out the series with few items, which introduce considerable noise. This is the reason they are not ranked on the web pages.

Of particular interest here is to compare the impact of journal articles relative to working papers. Table 3 shows that there is no clear winner, which could surprise many. We have to keep in mind that some journals have very low impact factors, while some working paper series have impact factors superior to most journals. Note also, as explained in the previous sections, that if the article version of a paper is cited, it counts towards both. So these numbers do not reflect where the citing author found the reference.

9.2 Works

How do the various rankings compare? Taking all articles and papers that are ranked in top 500 in any of the six categories on September 16, 2007, and narrowing them down to those listed in all six categories, we obtain a sample of 416 items. The fact that 83% of the top 500 according to one criterion are listed in all other criteria is already an indication of high correlation. Within this (rather small) sample, the rank correlations are still fairly high, averaging at 0.647 (Table 4). Rank correlations over the whole sample would be much larger, as demonstrated in other contexts below, but much more difficult to compute, for technical reasons.

9.3 Authors

We have 32 different ways to rank authors, thus if we want to compare how differently they perform, we need to look at 992 correlations ($32^2 - 32$). Table 5 reports them. While all these numbers can be overwhelming, the following can be extracted: The average correlation stands at 0.834 and varies between 0.608 and 0.998. The table groups the criteria in categories (number of works, citations, derived from citations, article pages, visibility on RePEc), and not surprisingly, correlations within these categories tend to be higher than with other categories. It is more interesting to see where criteria seem to differ most: article pages and visibility on RePEc, with an average correlation

of 0.660. This does not mean that they are orthogonal, though, 0.660 is still a significant correlation. But it is revealing that publishing in journals or even in good journals, has *relatively* little to do with how much people read ones work on RePEc.

Speaking of significance of correlations, there is a statistic that allows to measure how independent the criteria are from each other, χ^2. Here,

$$\chi^2_{n-1} = (n-1)((p-1)\bar{r}+1),$$

, where $p = 32$ is the number of criteria, $\bar{r} = 0.834$ is the average correlation, $n = 13946$ is the number of authors and $\chi^2_1 3945 = 374,479$. To be significant at 5%, the statistic would need to be below $14,211$. Therefore, we easily reject the null hypothesis that the criteria are independent.

Looking at only the 1000 top authors (Table 6, considering the 1000 authors with the most listed works), the correlations are smaller, between 0.185 and 0.998 averaging at 0.670, but follow the same patterns as above. The lowest correlation by criterion category is again between articles pages and RePEc visibility, at 0.281. While this seems a small number, one show take into account that this is within a sub-sample of authors that are jointly different from the rest of the sample (they all have a lot of publications). Again, if we apply the χ^2 statistics, we find $21,768$, which is much below the 5% threshold of $1,075$.

It should surprise no one that correlations are higher when when consider the aggregate ranking criteria, see Table 7. They average at 0.953, with a minimum of 0.861 and a maximum of 1. Excluding the best and worst criterion for each author does not make a significant impact on the overall picture, however, experience shows that it can alter the rankings at the very top for a few authors with a large variance in the rankings. The only exception in the "percent" aggregation, where are strong lead in a category can be devastating when it is excluded, for example. It is also remarkable that harmonic, arithmetic and geometric aggregation are all very close to each other.

As for individual criteria, correlations are lower when looking at the top 1000 authors, fluctuating between 0.379 and 1 for an average at 0.794, see Table 8. The patterns across aggregation criteria are similar to the full sample. For additional statistics for other sub-samples, see Table 9. Interestingly, in some small sub-samples for lower ranked authors, some correlations between individual criteria can get negative. Lower ranks are characterized by many ties (one or two citations, publications in series with a zero impact factors), and very little can mean large changes in rankings. But mean correlations are still high, despite these "accidents."

9.4 Institutions

The concordance of rankings across institutions is higher than that of authors for individual criteria, but the opposite is generally true for aggregate criteria, see Tables 10 to 14. We have no conceptual explanation for this difference at this point. Looking at the invidual correlations, the patterns are also somewhat differents to authors. For example, the h-index rankings typically correlate the least, while they were average for authors. Page counts for institutions correlate just as well with RePEc visibility as with other cirteria, while they were markedly lower for authors.

10 Comparison with Other Ranking Methodologies

The goal of this section is not to compare how the impact factors or rankings obtained by RePEc differ from other exercises.[7] It is rather to highlight some of the conceptual differences: what RePEc may miss and what others may miss.

10.1 What RePEc Can Do and Others Not

The rankings described above make use of the many facets of the data collected within the RePEc project. Some of them a quite unique, which certainly gives these rankings some added value when compared to existing rankings:

1. **Timeliness:** The data in RePEc is constantly updated and the results are continuously refreshed on its websites. For example, a working paper or article is typically listed within 24 hours of the publisher indexing it, its citation analysis is released within a month and its downloads are continuously monitored.

2. **Current affiliations:** Rankings of institutions reflect the current affiliations of authors and can take the move of an author from one to the other into account within a month. Other counts typically only take into account the affiliation at the time of publication.

3. **Pre-publications:** Established citation aggregators typically only consider citations in journals to journal articles. Even, the set of journals is often severely limited. There are no such restrictions in RePEc. In fact, working papers are a very important means of dissemination in Economics (and RePEc may have contributed to this) that should not be neglected. Note that analyzing working papers also significantly contributes to the timeliness of rankings.

4. **Certainty about authorship:** Given that authors acknowledge what works they have authored when they maintain their RePEc profiles, one big issue in ranking authors is resolved: name ambiguities. Indeed, many publications provide only the initial of the first name. Also, there are homonyms in the profession. RePEc data leaves no doubt.

5. **New ranking criteria:** Thanks to the fact that authors build profiles in RePEc, it is possible to reliably count how many different authors cite a particular author. We do not know of the use of the NCAuthors and RCAuthors criteria elsewhere.

10.2 What RePEc Cannot Do

There is very little human intervention in anything that RePEc does. Thus various aspects of other ranking analyses cannot be performed here:

1. **Errors:** Citation analysis is very much based on automatic reference extraction from texts and pattern matching of titles. Errors can obviously happen, and probably more so than with analysis by humans. The most important case is when a

[7] For a list of such ranking exercises, as indexed on RePEc, see http://ideas.repec.org/k/ranking.html

list of other working papers in a particular series is printed on the last page of a paper, and this list is interpreted as the continuation of the citations. Authors can now remove citations that are not accurate, though.

2. **Adjustments:** Any criteria based on page counts can be adjusted by the size of the page or its average word count in order to truly reflect the length of the article. RePEc does not do this, as it is completely automated.

3. **Stable impact factors:** Due to the constant adjustments in RePEc, impact factors change frequently, within bounds. But this makes the ouse of such factors difficult for third parties.

4. **Comprehensiveness:** Some important publications are still missing in RePEc, but RePEc has no staff to index them. Also, not all authors are registered with RePEc.

11 Conclusions

In this paper, we hope to have demonstrated that the ranking exercises performed in RePEc are based on a sound methodology and can be useful. It should also be clear that they are a work in progress, as the data is not yet as comprehensive as it could be, both in terms of listed publications and, especially, registered authors. The citation database is the component that is the most experimental at this point, as reference extraction and matching is difficult and error prone. As more publishers and more authors join in the RePEc project, as we perfect the analysis of the data, our confidence in the rankings will rise, and we hope the RePEc rankings will be regarded as a useful tool in the profession.

12 References

Sergey Brin and Larry Page, 1998. "The Anatomy of a Large-Scale Hypertextual Web Search Engine", *Computer Networks* 30(1-7): 107–117.

Jorge E. Hirsch, 2005. "An index to quantify an individual's scientific out", *Proceedings of the National Academy of Science*, 102: 16569, accessible at http://arxiv.org/abs/physics/0508025.

Thomas Krichel and Christian Zimmermann, 2005. "The Economics of Open Source Bibliographic Data Provision", University of Connecticut Department of Economics Working Paper 2005-01.

Frances Ruane and Richard S. J. Tol, 2007. "Rational (Successive) H-Indices: An Application to Economics in the Republic of Ireland", *Scientometrics*, forthcoming.

András Schubert, 2007. "Successive h-indices", *Scientometrics* 70(1): 201–205.

Table 1: Rank correlations of series

Impact factor	All				with ≥ 50 items			
	All series							
Simple factors	1	.974	.99	.958	1	.996	.999	.993
Recursive factors	.974	1	.969	.957	.996	1	.995	.991
Discounted factors	.99	.969	1	.964	.999	.995	1	.993
Recursive discounted factors	.958	.957	.964	1	.993	.991	.993	1
	Journals							
Simple factors	1	1	1	1	1	1	1	1
Recursive factors	1	1	1	1	1	1	1	.999
Discounted factors	1	1	1	1	1	1	1	.999
Recursive discounted factors	1	1	1	1	1	.999	.999	1
	Working paper series							
Simple factors	1	.99	.996	.984	1	.999	1	.999
Recursive factors	.99	1	.987	.983	.999	1	.999	.997
Discounted factors	.996	.987	1	.987	1	.999	1	.998
Recursive discounted factors	.984	.983	.987	1	.999	.997	.998	1

Table 2: Rank correlations of series (top 100 series in each pannel)

Impact factor	All				w/ ≥ 50 items			
	All series							
Simple factors	1	.223	1	.053	1	.944	.1	.677
Recursive factors	.223	1	.223	.74	.944	1	.944	.881
Discounted factors	1	.223	1	.053	1	.944	1	.677
Recursive discounted factors	.053	.74	.053	1	.677	.881	.667	1
	Journals							
Simple factors	1	1	1	1	1	.763	.935	.731
Recursive factors	1	1	1	1	.764	1	.721	.807
Discounted factors	1	1	1	1	.935	.721	1	.677
Recursive discounted factors	1	1	1	1	.731	.807	.677	1
	Working paper series							
Simple factors	1	.288	1	.194	1	.807	.893	.732
Recursive factors	.288	1	.288	.744	.807	1	.783	.622
Discounted factors	1	.288	1	.194	.893	.783	1	.751
Recursive discounted factors	.194	.744	.194	1	.732	.623	.751	1

Table 3: Average impact factors

	Papers	Journals
Simple factors	2.49	2.41
Recursive factors	0.41	0.22
Discounted simple factors	0.79	0.69
Discounted recursive factors	0.31	0.35

Table 4: Rank correlations of scores for top items by criteria

	Criteria from left column					
Number of citations	1	.709	.581	.718	.557	.492
Simple factors	.709	1	.932	.521	.702	.644
Recursive factors	.581	.932	1	.424	.647	.604
Discounted citations	.718	.521	.424	1	.668	.621
Discounted simple factors	.557	.702	.647	.668	1	.88
Discounted recursive factors	.492	.644	.604	.621	.88	1

Table 5: Rank correlations across criteria for authors, full sample

	Nb Works	DNb Works	Sc Works	WSc Works	ANb Works	ASc Works	AWSc Works	Nb Cites	D Cites	Sc Cites	DSc Cites	WSc Cites	WDSc Cites	ANb Cites	AD Cites	ASc Cites	ADSc Cites	AWSc Cites	AWDSc Cites	H Index	NC Authors	RC Authors	Nb Pages	Sc Pages	WSc Pages	ANb Pages	ASc Pages	AWSc Pages	Abs Views	Down loads	AAbs Views	ADown loads
NbWorks	1	.989	.875	.822	.96	.864	.814	.831	.813	.794	.785	.779	.784	.826	.811	.793	.784	.777	.782	.819	.832	.824	.845	.808	.791	.828	.8	.784	.902	.849	.865	.808
DNbWorks	.989	1	.848	.791	.969	.842	.786	.807	.787	.765	.754	.749	.753	.805	.787	.766	.754	.75	.753	.796	.808	.799	.849	.795	.776	.834	.789	.771	.889	.835	.866	.807
ScWorks	.875	.848	1	.979	.832	.989	.97	.871	.856	.879	.868	.873	.868	.868	.855	.878	.867	.872	.867	.853	.873	.875	.815	.886	.884	.8	.878	.878	.79	.748	.759	.713
WScWorks	.822	.791	.979	1	.778	.967	.991	.839	.828	.863	.854	.864	.853	.836	.827	.861	.852	.862	.852	.818	.843	.848	.76	.847	.854	.746	.839	.847	.742	.708	.709	.672
ANbWorks	.96	.969	.832	.778	1	.856	.799	.774	.755	.739	.728	.725	.727	.79	.772	.752	.74	.735	.739	.761	.774	.767	.831	.785	.768	.844	.795	.776	.861	.807	.889	.827
AScWorks	.864	.842	.989	.967	.856	1	.978	.849	.834	.859	.848	.854	.848	.859	.845	.867	.856	.861	.856	.831	.851	.853	.808	.878	.877	.81	.882	.881	.778	.735	.777	.728
AWScWorks	.814	.786	.97	.991	.799	.978	1	.822	.81	.847	.837	.849	.837	.829	.819	.853	.843	.854	.843	.8	.825	.831	.754	.842	.85	.754	.844	.852	.732	.697	.725	.686
NbCites	.831	.807	.871	.839	.774	.849	.822	1	.983	.974	.964	.958	.963	.991	.976	.968	.959	.953	.957	.957	.977	.976	.808	.847	.837	.783	.832	.824	.772	.744	.729	.698
DCites	.813	.787	.856	.828	.755	.834	.81	.983	1	.961	.976	.947	.974	.973	.992	.954	.97	.941	.968	.938	.964	.964	.788	.831	.822	.763	.815	.808	.761	.738	.716	.691
ScCites	.794	.765	.879	.863	.739	.859	.847	.974	.961	1	.988	.995	.988	.967	.955	.995	.984	.99	.983	.929	.963	.971	.781	.845	.843	.759	.832	.831	.738	.708	.696	.664
DScCites	.785	.754	.868	.854	.728	.848	.837	.964	.976	.988	1	.984	.998	.956	.97	.982	.995	.979	.993	.922	.957	.964	.766	.833	.83	.744	.819	.818	.733	.709	.689	.663
WScCites	.779	.749	.873	.864	.725	.854	.849	.958	.947	.995	.984	1	.983	.952	.943	.99	.98	.996	.979	.918	.95	.96	.767	.838	.839	.747	.826	.828	.724	.695	.683	.652
WDScCites	.784	.753	.868	.853	.727	.848	.837	.963	.974	.988	.998	.983	1	.955	.968	.982	.993	.978	.995	.921	.955	.963	.766	.833	.832	.745	.82	.82	.732	.707	.689	.662
ANbCites	.826	.805	.868	.836	.79	.859	.829	.991	.973	.967	.956	.952	.955	1	.982	.973	.963	.957	.961	.946	.967	.967	.812	.851	.841	.799	.844	.835	.766	.736	.741	.709
ADCites	.811	.787	.855	.827	.772	.845	.819	.976	.992	.955	.97	.943	.968	.982	1	.961	.976	.947	.974	.934	.957	.957	.793	.836	.827	.781	.829	.821	.757	.733	.729	.703
AScCites	.793	.766	.878	.861	.752	.867	.853	.968	.954	.995	.982	.99	.982	.973	.961	1	.988	.995	.988	.925	.956	.965	.784	.848	.846	.771	.841	.84	.734	.704	.705	.673
ADScCites	.784	.754	.867	.852	.74	.856	.843	.959	.97	.984	.995	.98	.993	.963	.976	.988	1	.984	.998	.919	.951	.958	.77	.836	.834	.757	.828	.827	.729	.704	.699	.671
AWScCites	.777	.75	.872	.862	.735	.861	.854	.953	.941	.99	.979	.996	.978	.957	.947	.995	.984	1	.983	.915	.944	.954	.77	.841	.841	.757	.834	.835	.72	.691	.69	.659
AWDScCites	.782	.753	.867	.852	.739	.856	.843	.957	.968	.983	.993	.979	.995	.961	.974	.988	.998	.983	1	.918	.949	.957	.77	.837	.835	.757	.829	.828	.728	.703	.698	.67
HIndex	.819	.796	.853	.818	.761	.831	.8	.957	.938	.929	.922	.918	.921	.946	.934	.925	.919	.915	.918	1	.934	.927	.784	.817	.808	.76	.804	.797	.754	.725	.708	.676
NCAuthors	.832	.808	.873	.843	.774	.851	.825	.977	.964	.963	.957	.95	.955	.967	.957	.956	.951	.944	.949	.934	1	.996	.796	.84	.831	.77	.825	.818	.781	.755	.735	.707
RCAuthors	.824	.799	.875	.848	.767	.853	.831	.976	.964	.971	.964	.96	.963	.967	.957	.965	.958	.954	.957	.927	.996	1	.797	.846	.838	.772	.831	.825	.772	.745	.727	.697
NbPages	.845	.849	.815	.76	.831	.808	.754	.808	.788	.781	.766	.767	.766	.812	.793	.784	.77	.77	.77	.784	.796	.797	1	.929	.907	.982	.924	.903	.719	.668	.703	.647
ScPages	.808	.795	.886	.847	.785	.878	.842	.847	.831	.845	.833	.838	.833	.851	.836	.848	.836	.841	.837	.817	.84	.846	.929	1	.994	.916	.993	.989	.693	.649	.671	.624
WScPages	.791	.776	.884	.854	.768	.877	.85	.837	.822	.843	.83	.839	.832	.841	.827	.846	.834	.841	.835	.808	.831	.838	.907	.994	1	.896	.988	.995	.677	.633	.656	.608
ANbPages	.828	.834	.8	.746	.844	.81	.754	.783	.763	.759	.744	.747	.745	.799	.781	.771	.757	.757	.757	.76	.77	.772	.982	.916	.896	1	.93	.907	.7	.648	.709	.651
AScPages	.8	.789	.878	.839	.795	.882	.844	.832	.815	.832	.819	.826	.82	.844	.829	.841	.828	.834	.829	.804	.825	.831	.924	.993	.988	.93	1	.994	.682	.638	.676	.627
AWScPages	.784	.771	.878	.847	.776	.881	.852	.824	.808	.831	.818	.828	.82	.835	.821	.84	.827	.835	.828	.797	.818	.825	.903	.989	.995	.907	.994	1	.668	.624	.659	.611
AbsViews	.902	.889	.79	.742	.861	.778	.732	.772	.761	.738	.733	.724	.732	.766	.757	.734	.729	.72	.728	.754	.781	.772	.719	.693	.677	.7	.682	.668	1	.957	.957	.913
Downloads	.849	.835	.748	.708	.807	.735	.697	.744	.738	.708	.709	.695	.707	.736	.733	.704	.704	.691	.703	.725	.755	.745	.668	.649	.633	.648	.638	.624	.957	1	.913	.957
AAbsViews	.865	.866	.759	.709	.889	.777	.725	.729	.716	.696	.689	.683	.689	.741	.729	.705	.699	.69	.698	.708	.735	.727	.703	.671	.656	.709	.676	.659	.957	.913	1	.95
ADownloads	.808	.807	.713	.672	.827	.728	.686	.698	.691	.664	.663	.652	.662	.709	.703	.673	.671	.659	.67	.676	.707	.697	.647	.624	.608	.651	.627	.611	.913	.957	.95	1

Table 6: Rank correlations across criteria for authors, top 1000 authors

	Nb Works	DNb Works	Sc Works	WSc Works	ANb Works	ASc Works	AWSc Works	Nb Cites	D Cites	Sc Cites	DSc Cites	WSc Cites	WDSc Cites	ANb Cites	AD Cites	ASc Cites	ADSc Cites	AWSc Cites	AWDSc Cites	H Index	NC Authors	RC Authors	Nb Pages	Sc Pages	WSc Pages	ANb Pages	ASc Pages	AWSc Pages	Abs Views	Down loads	AAbs Views	ADown loads
NbWorks	1	.841	.476	.432	.695	.458	.413	.468	.456	.444	.44	.43	.444	.457	.447	.434	.431	.42	.435	.434	.454	.452	.355	.378	.373	.319	.359	.355	.636	.516	.539	.444
DNbWorks	.841	1	.271	.226	.765	.274	.225	.247	.228	.216	.201	.208	.204	.248	.23	.216	.202	.209	.204	.209	.239	.236	.304	.222	.213	.299	.22	.211	.458	.37	.466	.377
ScWorks	.476	.271	1	.952	.347	.98	.934	.833	.801	.875	.846	.869	.85	.833	.802	.871	.844	.862	.848	.815	.841	.854	.728	.903	.891	.678	.884	.873	.425	.396	.374	.364
WScWorks	.432	.226	.952	1	.344	.946	.988	.768	.751	.85	.829	.865	.834	.775	.756	.849	.828	.86	.833	.757	.785	.804	.705	.89	.899	.679	.877	.885	.374	.35	.346	.333
ANbWorks	.695	.765	.347	.344	1	.445	.412	.275	.268	.295	.282	.307	.285	.345	.333	.346	.331	.35	.332	.245	.272	.281	.364	.342	.347	.493	.404	.401	.373	.296	.583	.447
AScWorks	.458	.274	.98	.946	.445	1	.957	.79	.763	.846	.818	.848	.822	.816	.787	.862	.834	.859	.838	.772	.8	.816	.724	.901	.894	.721	.91	.9	.393	.363	.412	.385
AWScWorks	.413	.225	.934	.988	.412	.957	1	.732	.717	.823	.801	.843	.806	.758	.74	.837	.815	.85	.819	.722	.751	.772	.699	.884	.896	.706	.892	.901	.348	.322	.371	.345
NbCites	.468	.247	.833	.768	.275	.79	.732	1	.971	.959	.945	.928	.941	.987	.959	.942	.932	.912	.929	.976	.975	.97	.608	.79	.77	.541	.754	.738	.515	.519	.428	.45
DCites	.456	.228	.801	.751	.268	.763	.717	.971	1	.934	.962	.907	.956	.958	.988	.918	.949	.891	.943	.95	.952	.946	.591	.769	.749	.532	.734	.716	.532	.539	.443	.469
ScCites	.444	.216	.875	.85	.295	.846	.823	.959	.934	1	.981	.991	.981	.956	.931	.991	.974	.98	.975	.943	.962	.971	.653	.851	.85	.608	.825	.824	.46	.454	.408	.417
DScCites	.44	.201	.846	.829	.282	.818	.801	.945	.962	.981	1	.972	.998	.94	.958	.971	.992	.96	.989	.932	.951	.958	.631	.828	.825	.589	.8	.798	.487	.481	.427	.439
WScCites	.43	.208	.869	.865	.307	.848	.843	.928	.907	.991	.972	1	.972	.931	.909	.986	.969	.993	.97	.914	.935	.948	.653	.854	.86	.621	.834	.84	.437	.428	.402	.404
WDScCites	.444	.204	.85	.834	.285	.822	.806	.941	.956	.981	.998	.972	1	.937	.951	.971	.989	.961	.991	.928	.949	.957	.632	.831	.83	.59	.803	.803	.486	.479	.425	.437
ANbCites	.457	.248	.833	.775	.345	.816	.758	.987	.958	.956	.94	.931	.937	1	.971	.96	.946	.932	.943	.966	.962	.96	.62	.806	.788	.587	.789	.772	.496	.497	.464	.472
ADCites	.447	.23	.802	.756	.333	.787	.74	.959	.988	.931	.958	.909	.951	.971	1	.934	.963	.909	.957	.942	.941	.937	.601	.781	.763	.572	.764	.746	.517	.523	.479	.494
AScCites	.434	.216	.871	.849	.346	.862	.837	.942	.918	.991	.971	.986	.971	.96	.934	1	.981	.991	.981	.929	.947	.959	.658	.857	.857	.638	.846	.845	.446	.438	.435	.435
ADScCites	.431	.202	.844	.828	.331	.834	.815	.932	.949	.974	.992	.969	.989	.946	.963	.981	1	.972	.998	.922	.939	.949	.635	.833	.832	.617	.82	.818	.474	.468	.454	.459
AWScCites	.42	.209	.862	.86	.35	.859	.85	.912	.891	.98	.96	.993	.961	.932	.909	.991	.972	1	.973	.9	.92	.935	.654	.856	.863	.643	.849	.854	.425	.415	.425	.42
AWDScCites	.435	.204	.848	.833	.332	.838	.819	.929	.943	.975	.989	.97	.991	.943	.957	.981	.998	.973	1	.918	.938	.948	.636	.837	.837	.617	.823	.823	.472	.465	.45	.455
HIndex	.434	.209	.815	.757	.245	.772	.722	.976	.95	.943	.932	.914	.928	.966	.942	.929	.922	.9	.918	1	.955	.952	.596	.777	.76	.531	.742	.728	.481	.479	.394	.411
NCAuthors	.454	.239	.841	.785	.272	.8	.751	.975	.952	.962	.951	.935	.949	.962	.941	.947	.939	.92	.938	.955	1	.997	.629	.806	.787	.566	.772	.756	.519	.529	.438	.466
RCAuthors	.452	.236	.854	.804	.281	.816	.772	.97	.946	.971	.958	.948	.957	.96	.937	.959	.949	.935	.948	.952	.997	1	.643	.824	.809	.586	.792	.779	.507	.512	.433	.456
NbPages	.355	.304	.728	.705	.364	.724	.699	.608	.591	.653	.631	.653	.632	.62	.601	.658	.635	.654	.636	.596	.629	.643	1	.833	.782	.958	.817	.768	.23	.231	.266	.263
ScPages	.378	.222	.903	.89	.342	.901	.884	.79	.769	.851	.828	.854	.831	.806	.781	.857	.833	.856	.837	.777	.806	.824	.833	1	.987	.81	.987	.973	.308	.304	.302	.307
WScPages	.373	.213	.891	.899	.347	.894	.896	.77	.749	.85	.825	.86	.83	.788	.763	.857	.832	.863	.837	.76	.787	.809	.782	.987	1	.771	.979	.99	.298	.29	.3	.3
ANbPages	.319	.299	.678	.679	.493	.721	.706	.541	.532	.608	.589	.621	.59	.587	.572	.638	.617	.643	.617	.531	.566	.586	.958	.81	.771	1	.833	.79	.187	.185	.319	.292
AScPages	.359	.22	.884	.877	.404	.91	.892	.754	.734	.825	.8	.834	.803	.789	.764	.846	.82	.849	.823	.742	.772	.792	.817	.987	.979	.833	1	.988	.281	.276	.325	.318
AWScPages	.355	.211	.873	.885	.401	.9	.901	.738	.716	.824	.798	.84	.803	.772	.746	.845	.818	.854	.823	.728	.756	.779	.768	.973	.99	.79	.988	1	.273	.265	.318	.309
AbsViews	.636	.458	.425	.374	.373	.393	.348	.515	.532	.46	.487	.437	.486	.496	.517	.446	.474	.425	.472	.481	.519	.507	.23	.308	.298	.187	.281	.273	1	.899	.877	.804
Downloads	.516	.37	.396	.35	.296	.363	.322	.519	.539	.454	.481	.428	.479	.497	.523	.438	.468	.415	.465	.479	.529	.512	.231	.304	.29	.185	.276	.265	.899	1	.785	.913
AAbsViews	.539	.466	.374	.346	.583	.412	.371	.428	.443	.408	.427	.402	.425	.464	.479	.435	.454	.425	.45	.394	.438	.433	.266	.302	.3	.319	.325	.318	.877	.785	1	.876
ADownloads	.444	.377	.364	.333	.447	.385	.345	.45	.469	.417	.439	.404	.437	.472	.494	.435	.459	.42	.455	.411	.466	.456	.263	.307	.3	.292	.318	.309	.804	.913	.876	1

Table 7: Rank correlations across aggregate criteria for authors, full sample

		harmonic		arithmetic		geometric		lexicographic		graphicolexic		percent	
exclude outliers?		no	yes	no	yes	no	yes	no	yes	no	yes	no	yes
harmonic	no	1	1	.9909	.9911	.996	.9961	.9684	.9681	.9189	.9189	.9897	.9195
harmonic	yes	1	1	.9908	.991	.9959	.996	.9686	.9683	.918	.918	.9896	.9186
arithmetic	no	.9909	.9908	1	1	.9979	.9978	.9388	.9384	.9417	.9417	.9821	.942
arithmetic	yes	.9911	.991	1	1	.998	.9979	.9392	.9388	.9398	.9398	.9821	.9401
geometric	no	.996	.9959	.9979	.998	1	1	.9522	.9518	.9313	.9313	.9861	.9318
geometric	yes	.9961	.996	.9978	.9979	1	1	.9525	.9521	.93	.93	.9861	.9304
lexicographic	no	.9684	.9686	.9388	.9392	.9522	.9525	1	1	.8616	.8616	.9683	.8626
lexicographic	yes	.9681	.9683	.9384	.9388	.9518	.9521	1	1	.8607	.8607	.9679	.8617
graphicolexic	no	.9189	.918	.9417	.9398	.9313	.93	.8616	.8607	1	1	.9272	1
graphicolexic	yes	.9189	.918	.9417	.9398	.9313	.93	.8616	.8607	1	1	.9272	1
percent	no	.9897	.9896	.9821	.9821	.9861	.9861	.9683	.9679	.9272	.9272	1	.9278
percent	yes	.9195	.9186	.942	.9401	.9318	.9304	.8626	.8617	1	1	.9278	1

Table 8: Rank correlations across aggregate criteria for authors, top 1000 authors

		harmonic		arithmetic		geometric		lexicographic		graphicolexic		percent	
exclude outliers?		no	yes	no	yes	no	yes	no	yes	no	yes	no	yes
harmonic	no	1	1	.8118	.8139	.9179	.9195	.8389	.8388	.6706	.6706	.9827	.6711
harmonic	yes	1	1	.8109	.813	.9172	.9189	.8399	.8398	.6692	.6692	.9826	.6697
arithmetic	no	.8118	.8109	1	1	.9734	.9724	.5081	.5079	.9369	.9369	.8743	.9369
arithmetic	yes	.8139	.813	1	1	.9743	.9733	.5106	.5105	.9342	.9342	.8761	.9341
geometric	no	.9179	.9172	.9734	.9743	1	1	.6421	.642	.8734	.8734	.9556	.8736
geometric	yes	.9195	.9189	.9724	.9733	1	1	.6447	.6445	.8708	.8708	.9567	.871
lexicographic	no	.8389	.8399	.5081	.5106	.6421	.6447	1	1	.379	.379	.7998	.3796
lexicographic	yes	.8387	.8398	.5079	.5105	.642	.6445	1	1	.3788	.3788	.7997	.3795
graphicolexic	no	.6706	.6692	.9369	.9342	.8734	.8708	.379	.3788	1	1	.7428	1
graphicolexic	yes	.6706	.6692	.9369	.9342	.8734	.8708	.379	.3788	1	1	.7428	1
percent	no	.9827	.9826	.8743	.8761	.9556	.9567	.7998	.7997	.7428	.7428	1	.7433
percent	yes	.6711	.6697	.9369	.9341	.8736	.871	.3796	.3795	1	1	.7433	1

Table 9: Average correlations across criteria for authors

	Individual criteria			Aggregate criteria		
Sample	mean	max	min	mean	max	min
Full	.834	.998	.608	.953	1	.861
1–250	.636	.998	-.065	.670	1	.096
1–500	.641	.998	.060	.748	1	.283
1–750	.659	.998	.160	.768	1	.304
1–1000	.670	.998	.185	.794	1	.379
1–2000	.681	.997	.216	.831	1	.476
1–3000	.694	.997	.253	.857	1	.563
1–4000	.704	.997	.295	.871	1	.605
4001–8000	.562	.996	-.061	.834	1	.507
8001–12000	.469	.995	-.107	.806	1	.474
1001–2000	.603	.997	-.040	.810	1	.391
2001–3000	.578	.997	-.066	.813	1	.434
3001–4000	.552	.996	-.076	.809	1	.408
4001–5000	.529	.996	-.182	.837	1	.488
5001–6000	.481	.995	-.236	.841	1	.521
6001–7000	.475	.995	-.222	.813	1	.458

Table 10: Rank correlations across criteria for institutions, full sample

	Nb Works	DNb Works	Sc Works	WSc Works	ANb Works	ASc Works	AWSc Works	Nb Cites	D Cites	Sc Cites	DSc Cites	WSc Cites	WDSc Cites	ANb Cites	AD Cites	ASc Cites	ADSc Cites	AWSc Cites	AWDSc Cites	H Index	NC Authors	RC Authors	Nb Pages	Sc Pages	WSc Pages	ANb Pages	ASc Pages	AWSc Pages	Abs Views	Down loads	AAbs Views	ADown loads
NbWorks	1	.995	.88	.819	.99	.863	.803	.775	.797	.717	.744	.692	.741	.756	.781	.701	.729	.677	.727	.751	.835	.828	.964	.85	.821	.954	.836	.808	.949	.923	.926	.891
DNbWorks	.995	1	.856	.791	.992	.839	.776	.747	.769	.686	.713	.661	.71	.729	.754	.672	.7	.647	.697	.753	.812	.805	.964	.829	.797	.954	.815	.785	.943	.917	.924	.888
ScWorks	.88	.856	1	.986	.889	.996	.98	.959	.963	.936	.946	.922	.945	.951	.959	.927	.94	.914	.94	.695	.978	.978	.916	.985	.98	.922	.981	.976	.914	.921	.918	.918
WScWorks	.819	.791	.986	1	.832	.986	.996	.964	.963	.956	.959	.949	.958	.958	.961	.947	.954	.94	.954	.657	.968	.969	.86	.968	.971	.869	.967	.97	.874	.887	.884	.892
ANbWorks	.99	.992	.889	.832	1	.88	.823	.784	.803	.729	.752	.704	.749	.772	.794	.719	.745	.696	.742	.752	.843	.838	.973	.862	.834	.972	.855	.828	.946	.923	.938	.905
AScWorks	.863	.839	.996	.986	.88	1	.987	.956	.958	.935	.942	.922	.94	.955	.961	.933	.944	.92	.943	.682	.973	.973	.904	.981	.976	.917	.984	.979	.899	.907	.911	.912
AWScWorks	.803	.776	.98	.996	.823	.987	1	.958	.954	.95	.951	.944	.95	.959	.959	.949	.953	.942	.953	.643	.959	.961	.847	.961	.964	.862	.966	.969	.858	.872	.876	.885
NbCites	.775	.747	.959	.964	.784	.956	.958	1	.998	.993	.994	.986	.993	.993	.994	.985	.99	.978	.989	.622	.986	.987	.824	.956	.963	.831	.953	.959	.865	.894	.877	.901
DCites	.797	.769	.963	.963	.803	.958	.954	.998	1	.986	.993	.978	.992	.987	.993	.975	.985	.967	.984	.638	.989	.989	.84	.958	.962	.844	.952	.956	.882	.91	.889	.911
ScCites	.717	.686	.936	.956	.729	.935	.95	.993	.986	1	.997	.998	.997	.989	.986	.994	.994	.992	.994	.586	.969	.971	.771	.935	.95	.78	.933	.947	.821	.852	.837	.865
DScCites	.744	.713	.946	.959	.752	.942	.951	.994	.993	.997	1	.993	1	.987	.989	.987	.993	.984	.993	.605	.977	.978	.792	.942	.955	.799	.938	.949	.841	.871	.852	.878
WScCites	.692	.661	.922	.949	.704	.922	.944	.986	.978	.998	.993	1	.993	.982	.977	.992	.99	.994	.99	.569	.957	.96	.747	.921	.94	.756	.92	.937	.801	.834	.82	.85
WDScCites	.741	.71	.945	.958	.749	.94	.95	.993	.992	.997	1	.993	1	.985	.987	.987	.993	.983	.993	.604	.975	.977	.79	.941	.955	.796	.937	.949	.838	.868	.849	.875
ANbCites	.756	.729	.951	.958	.772	.955	.959	.993	.987	.989	.987	.982	.985	1	.998	.993	.995	.987	.994	.612	.981	.982	.81	.95	.958	.823	.954	.961	.849	.878	.872	.898
ADCites	.781	.754	.959	.961	.794	.961	.959	.994	.993	.986	.989	.977	.987	.998	1	.988	.994	.979	.993	.629	.988	.988	.83	.956	.962	.84	.957	.962	.869	.896	.887	.91
AScCites	.701	.672	.927	.947	.719	.933	.949	.985	.975	.994	.987	.992	.987	.993	.988	1	.997	.998	.997	.577	.963	.965	.759	.928	.944	.773	.933	.948	.807	.838	.834	.863
ADScCites	.729	.7	.94	.954	.745	.944	.953	.99	.985	.994	.993	.99	.993	.995	.994	.997	1	.994	1	.597	.974	.976	.783	.939	.952	.795	.941	.953	.829	.859	.851	.878
AWScCites	.677	.647	.914	.94	.696	.92	.942	.978	.967	.992	.984	.994	.983	.987	.979	.998	.994	1	.993	.561	.952	.955	.736	.915	.934	.75	.92	.937	.789	.821	.818	.849
AWDScCites	.727	.697	.94	.954	.742	.943	.953	.989	.984	.994	.993	.99	.993	.994	.993	.997	1	.993	1	.597	.973	.975	.781	.938	.952	.793	.941	.953	.827	.857	.849	.876
HIndex	.75	.753	.695	.657	.752	.682	.643	.622	.638	.586	.605	.569	.604	.612	.629	.576	.597	.561	.597	1	.677	.674	.758	.686	.67	.75	.674	.66	.724	.71	.716	.695
NCAuthors	.835	.812	.978	.968	.843	.973	.959	.986	.989	.969	.977	.957	.975	.981	.988	.963	.974	.952	.973	.677	1	1	.878	.972	.973	.883	.967	.967	.904	.925	.912	.928
RCAuthors	.828	.805	.978	.969	.838	.973	.961	.987	.989	.971	.978	.96	.977	.982	.988	.965	.976	.955	.975	.674	1	1	.874	.972	.974	.879	.967	.969	.898	.92	.908	.924
NbPages	.964	.964	.916	.86	.973	.904	.847	.824	.84	.771	.792	.747	.79	.81	.83	.759	.783	.736	.781	.758	.878	.874	1	.914	.887	.996	.904	.878	.934	.921	.927	.903
ScPages	.85	.829	.985	.968	.862	.981	.961	.956	.958	.935	.942	.921	.941	.95	.956	.928	.939	.915	.938	.686	.972	.972	.914	1	.996	.919	.997	.993	.891	.899	.899	.902
WScPages	.821	.797	.98	.971	.834	.976	.964	.963	.962	.95	.955	.94	.955	.958	.962	.944	.952	.934	.952	.67	.973	.974	.887	.996	1	.894	.993	.997	.872	.884	.884	.89
ANbPages	.954	.954	.922	.869	.972	.917	.862	.831	.844	.78	.799	.756	.796	.823	.84	.773	.795	.75	.793	.75	.883	.879	.996	.919	.894	1	.915	.891	.927	.914	.928	.904
AScPages	.836	.815	.981	.967	.855	.984	.966	.953	.952	.933	.938	.92	.937	.954	.957	.933	.941	.92	.941	.674	.967	.967	.904	.997	.993	.915	1	.996	.878	.887	.894	.897
AWScPages	.808	.785	.976	.97	.828	.979	.969	.959	.956	.947	.949	.937	.949	.961	.962	.948	.953	.937	.953	.66	.967	.969	.878	.993	.997	.891	.996	1	.86	.872	.879	.886
AbsViews	.949	.943	.914	.874	.946	.899	.858	.865	.882	.821	.841	.801	.838	.849	.869	.807	.829	.789	.827	.724	.904	.898	.934	.891	.872	.927	.878	.86	1	.989	.987	.972
Downloads	.923	.917	.921	.887	.923	.907	.872	.894	.91	.852	.871	.834	.868	.878	.896	.838	.859	.821	.857	.71	.925	.92	.921	.899	.884	.914	.887	.872	.989	1	.984	.987
AAbsViews	.926	.924	.918	.884	.938	.911	.876	.877	.889	.837	.852	.82	.849	.872	.887	.834	.851	.818	.849	.716	.912	.908	.927	.899	.884	.928	.894	.879	.987	.984	1	.99
ADownloads	.891	.888	.918	.892	.905	.912	.885	.901	.911	.865	.878	.85	.875	.898	.91	.863	.878	.849	.876	.695	.928	.924	.903	.902	.89	.904	.897	.886	.972	.987	.99	1

Table 11: Rank correlations across criteria for institutions, top 250 institutions

	Nb Works	DNb Works	Sc Works	WSc Works	ANb Works	ASc Works	AWSc Works	Nb Cites	D Cites	Sc Cites	DSc Cites	WSc Cites	WDSc Cites	ANb Cites	AD Cites	ASc Cites	ADSc Cites	AWSc Cites	AWDSc Cites	H Index	NC Authors	RC Authors	Nb Pages	Sc Pages	WSc Pages	ANb Pages	ASc Pages	AWSc Pages	Abs Views	Down loads	AAbs Views	ADown loads
NbWorks	1	.987	.794	.685	.976	.776	.669	.614	.634	.536	.556	.511	.55	.606	.63	.528	.553	.503	.547	.685	.713	.701	.901	.719	.663	.886	.709	.657	.901	.845	.833	.775
DNbWorks	.987	1	.753	.642	.984	.739	.629	.56	.577	.48	.495	.457	.488	.557	.578	.477	.496	.454	.49	.689	.667	.654	.909	.69	.632	.897	.686	.63	.883	.823	.832	.771
ScWorks	.794	.753	1	.969	.78	.988	.955	.857	.849	.833	.826	.821	.821	.859	.858	.831	.833	.819	.828	.537	.913	.915	.834	.937	.921	.826	.929	.915	.768	.758	.711	.696
WScWorks	.684	.642	.969	1	.674	.959	.988	.845	.827	.845	.826	.845	.82	.852	.841	.846	.836	.846	.831	.453	.878	.884	.73	.887	.889	.726	.88	.883	.689	.695	.642	.644
ANbWorks	.976	.984	.78	.675	1	.783	.676	.579	.594	.505	.519	.483	.512	.589	.608	.513	.531	.49	.525	.677	.683	.672	.908	.712	.657	.917	.722	.668	.876	.819	.85	.789
AScWorks	.776	.739	.988	.959	.783	1	.967	.828	.818	.807	.796	.795	.791	.851	.848	.824	.823	.812	.818	.519	.888	.891	.819	.919	.903	.83	.931	.916	.741	.73	.706	.689
AWScWorks	.667	.629	.955	.988	.676	.967	1	.814	.793	.815	.792	.815	.786	.839	.826	.834	.821	.834	.816	.435	.849	.856	.714	.864	.864	.724	.875	.876	.662	.666	.633	.634
NbCites	.614	.56	.857	.845	.579	.828	.814	1	.992	.985	.984	.973	.982	.979	.977	.96	.966	.948	.965	.435	.961	.959	.669	.864	.867	.651	.837	.84	.693	.737	.618	.657
DCites	.634	.577	.849	.827	.594	.818	.793	.992	1	.969	.985	.953	.982	.965	.978	.938	.961	.922	.959	.452	.956	.953	.677	.855	.853	.657	.825	.824	.713	.756	.63	.665
ScCites	.536	.48	.833	.845	.505	.807	.815	.985	.969	1	.99	.997	.99	.968	.958	.979	.977	.975	.977	.381	.936	.94	.605	.843	.859	.59	.817	.833	.624	.672	.557	.599
DScCites	.556	.495	.826	.826	.519	.796	.793	.984	.985	.99	1	.981	.999	.959	.964	.961	.977	.951	.978	.398	.937	.939	.613	.836	.847	.595	.807	.819	.645	.693	.567	.607
WScCites	.511	.457	.821	.845	.483	.795	.815	.973	.953	.997	.981	1	.981	.957	.943	.976	.968	.978	.969	.364	.921	.927	.584	.83	.85	.57	.804	.823	.603	.653	.54	.585
WDScCites	.55	.488	.821	.82	.512	.791	.786	.982	.982	.99	.999	.981	1	.956	.961	.959	.976	.95	.978	.394	.934	.936	.608	.833	.846	.591	.804	.816	.639	.688	.561	.602
ANbCites	.606	.557	.859	.852	.589	.851	.839	.979	.965	.968	.959	.957	.956	1	.992	.984	.983	.973	.981	.428	.951	.951	.669	.866	.869	.665	.857	.859	.678	.719	.627	.665
ADCites	.63	.578	.858	.841	.608	.848	.826	.977	.978	.958	.964	.943	.961	.992	1	.968	.984	.953	.981	.448	.952	.95	.682	.862	.861	.676	.851	.849	.702	.743	.643	.679
AScCites	.528	.477	.831	.846	.513	.824	.834	.96	.938	.979	.961	.976	.959	.984	.968	1	.99	.997	.989	.373	.923	.928	.602	.838	.855	.6	.831	.846	.607	.652	.563	.605
ADScCites	.553	.496	.833	.836	.531	.823	.821	.966	.961	.977	.977	.968	.976	.983	.984	.99	1	.982	.999	.393	.932	.935	.615	.839	.851	.611	.828	.839	.633	.679	.578	.619
AWScCites	.503	.454	.819	.846	.49	.812	.834	.948	.922	.975	.951	.978	.95	.973	.953	.997	.982	1	.981	.356	.908	.915	.581	.824	.845	.579	.817	.836	.586	.634	.545	.591
AWDScCites	.547	.49	.828	.831	.525	.818	.816	.965	.959	.977	.978	.969	.978	.981	.981	.989	.999	.981	1	.389	.929	.933	.61	.837	.85	.606	.825	.837	.627	.674	.572	.614
HIndex	.685	.689	.537	.453	.676	.519	.435	.435	.452	.381	.398	.364	.394	.428	.448	.373	.393	.356	.389	1	.526	.518	.658	.514	.478	.646	.507	.473	.63	.595	.587	.548
NCAuthors	.713	.667	.913	.878	.683	.888	.849	.961	.956	.936	.937	.921	.934	.951	.952	.923	.932	.908	.929	.526	1	.999	.764	.907	.9	.747	.887	.88	.759	.78	.687	.703
RCAuthors	.701	.654	.915	.884	.672	.891	.856	.959	.953	.94	.939	.927	.936	.951	.95	.928	.935	.915	.933	.518	.999	1	.756	.909	.904	.74	.89	.886	.746	.767	.676	.691
NbPages	.901	.909	.834	.73	.908	.819	.714	.669	.677	.605	.613	.584	.608	.669	.682	.602	.615	.581	.61	.658	.764	.756	1	.848	.798	.987	.842	.795	.805	.765	.758	.711
ScPages	.719	.69	.937	.887	.712	.919	.864	.864	.855	.843	.836	.83	.833	.866	.862	.838	.839	.824	.837	.514	.907	.909	.848	1	.992	.835	.988	.981	.706	.704	.657	.65
WScPages	.663	.632	.921	.889	.657	.903	.864	.867	.853	.859	.847	.85	.846	.869	.861	.855	.851	.845	.85	.478	.9	.904	.798	.992	1	.786	.979	.988	.667	.671	.621	.62
ANbPages	.886	.897	.826	.726	.917	.83	.724	.651	.657	.59	.595	.57	.591	.665	.676	.6	.611	.579	.606	.647	.747	.74	.987	.835	.787	1	.848	.801	.785	.743	.757	.708
AScPages	.709	.686	.929	.88	.722	.931	.875	.837	.825	.817	.807	.804	.804	.857	.851	.831	.828	.817	.825	.507	.887	.89	.842	.988	.979	.848	1	.992	.688	.682	.656	.645
AWScPages	.657	.63	.915	.883	.668	.916	.876	.84	.824	.833	.819	.823	.816	.859	.849	.846	.839	.836	.837	.473	.88	.886	.795	.981	.988	.801	.992	1	.651	.651	.621	.616
AbsViews	.901	.883	.768	.689	.876	.741	.662	.693	.713	.624	.645	.603	.639	.678	.702	.607	.633	.586	.627	.63	.759	.746	.805	.706	.667	.785	.688	.651	1	.973	.955	.928
Downloads	.845	.823	.758	.695	.819	.73	.666	.737	.756	.672	.693	.653	.688	.719	.743	.652	.679	.634	.674	.595	.78	.767	.765	.704	.671	.743	.682	.651	.973	1	.932	.956
AAbsViews	.833	.832	.711	.642	.85	.706	.633	.618	.63	.557	.567	.54	.561	.627	.643	.563	.578	.545	.572	.587	.687	.676	.758	.657	.621	.757	.656	.621	.955	.932	1	.972
ADownloads	.775	.771	.696	.644	.789	.689	.634	.657	.665	.599	.607	.585	.602	.665	.679	.605	.619	.591	.614	.548	.703	.691	.711	.65	.62	.708	.645	.616	.928	.956	.972	1

Table 12: Rank correlations across aggregate criteria for institutions, full sample

		harmonic		arithmetic		geometric		lexicographic		graphicolexic		percent	
exclude outliers?		no	yes	no	yes	no	yes	no	yes	no	yes	no	yes
harmonic	no	1	.9991	.614	.617	.7063	.7034	.843	.843	.504	.504	.5017	.5237
harmonic	yes	.9991	1	.6342	.6373	.7277	.7251	.8581	.8581	.5233	.5233	.5231	.5455
arithmetic	no	.614	.6342	1	.9996	.9391	.9367	.8047	.8047	.8947	.8947	.9293	.945
arithmetic	yes	.617	.6373	.9996	1	.9441	.9418	.8098	.8098	.8911	.8911	.9309	.9468
geometric	no	.7063	.7277	.9391	.9441	1	.9998	.9115	.9115	.8157	.8157	.8901	.9076
geometric	yes	.7034	.7251	.9367	.9418	.9998	1	.911	.911	.8127	.8127	.8899	.9074
lexicographic	no	.843	.8581	.8047	.8098	.9115	.911	1	1	.684	.684	.7417	.7606
lexicographic	yes	.843	.8581	.8047	.8098	.9115	.911	1	1	.684	.684	.7417	.7606
graphicolexic	no	.504	.5233	.8947	.8911	.8157	.8127	.684	.684	1	1	.8182	.8322
graphicolexic	yes	.504	.5233	.8947	.8911	.8157	.8127	.684	.684	1	1	.8182	.8322
percent	no	.5017	.5231	.9293	.9309	.8901	.8899	.7417	.7417	.8182	.8182	1	.9968
percent	yes	.5237	.5455	.945	.9468	.9076	.9074	.7606	.7606	.8322	.8322	.9968	1

Table 13: Rank correlations across aggregate criteria for institutions, top 250 institutions

		harmonic		arithmetic		geometric		lexicographic		graphicolexic		percent	
exclude outliers?		no	yes	no	yes	no	yes	no	yes	no	yes	no	yes
harmonic	no	1	.9999	.6102	.6072	.8119	.813	.9173	.9173	.6927	.6927	.3987	.4047
harmonic	yes	.9999	1	.6087	.6057	.8107	.8119	.918	.918	.6911	.6911	.3958	.4018
arithmetic	no	.6102	.6087	1	.9987	.7946	.7805	.5681	.5681	.9797	.9797	.6972	.7056
arithmetic	yes	.6072	.6057	.9987	1	.8016	.7883	.5689	.5689	.9709	.9709	.709	.7175
geometric	no	.8119	.8107	.7946	.8016	1	.9992	.8377	.8377	.798	.798	.6187	.6268
geometric	yes	.813	.8119	.7805	.7883	.9992	1	.8429	.8429	.7842	.7842	.6202	.6281
lexicographic	no	.9173	.918	.5681	.5689	.8377	.8429	1	1	.6293	.6293	.4246	.4303
lexicographic	yes	.9173	.918	.5681	.5689	.8377	.8429	1	1	.6293	.6293	.4246	.4303
graphicolexic	no	.6927	.6911	.9797	.9709	.798	.7842	.6293	.6293	1	1	.6773	.6853
graphicolexic	yes	.6927	.6911	.9797	.9709	.798	.7842	.6293	.6293	1	1	.6773	.6853
percent	no	.3987	.3958	.6972	.709	.6187	.6202	.4246	.4246	.6773	.6773	1	.9992
percent	yes	.4047	.4018	.7056	.7175	.6268	.6281	.4303	.4303	.6853	.6853	.9992	1

Table 14: Average correlations across criteria for institutions

	Individual criteria			Aggregate criteria		
Sample	mean	max	min	mean	max	min
Full	.892	1	.561	.799	.947	.502
1–250	.772	.999	.356	.725	.980	.396
1–500	.816	.999	.449	.751	.985	.429
1–750	.849	1	.492	.750	.985	.387
1–1000	.868	1	.528	.742	.986	.349
1001–2000	.840	1	.221	.926	.994	.790
2001–3000	.869	1	.269	.997	1	.987