



International Journal of HRM and Organizational Behavior



www.ijhrmob.com

editor@ijhrmob.com

An Alternative Metric for Assessing Journal Impact in Economics and Business: The Google Scholar H-Index Kapileshwar

Indian Institute of Management Calcutta

ABSTRACT

To better measure the significance of economics and business journals, we suggest using Google Scholar as a data source and Hirsch's h-index as a metric. Using a sample of 838 journals in Economics & Business, we find that the Google Scholar h-index provides a more accurate and complete estimate of journal impact than the ISI Journal Impact Factor.

INTRODUCTION

Although the creation of rankings of academic journals is common practice, the activity is not without contention or critique (cf. McDonald & Kam, 2007). Whilst recognising and sympathising with this position, the present paper takes a pragmatic stance: as long as journal rankings are considered to be part of academic life, it is important to ensure that they are as comprehensive and objective as possible. We can distinguish two approaches to ranking journals: stated preference and revealed preference (Tahai & Meyer, 1999). Stated preference involves members of a particular academic community ranking journals on the basis of their own expert judgements. There are hundreds of individual university journal rankings and collated journal ranking lists have sprung up (cf. the ABS Journal Quality Guide (ABS, 2007) and Harzing's Journal Quality List (Harzing, 2007). Rankings might be based on anything from a large-scale worldwide survey of academics to a small group of individuals with decision-making power, but will always contain some element of subjectivity. Revealed preference rankings are based on *actual* publication behaviour and generally measure the citation rates of journals using ISI's Journal Impact Factors (JIFs). The JIF is defined as the mean number of citations received in a particular year to articles published in the journal in the preceding two years. As the selection of article titles in Table 1 shows, this statistic is by no means undisputed.

Mingers & Harzing (2007) report a high degree of correlation between journal rankings based on stated and revealed preference. However, stated preference studies have long memories and perceptions of journals normally change very slowly (Tahai & Meyer, 1999). Revealed preference studies therefore provide a fairer assessment of new journals or journals that have recently improved their standing and are argued to present a more accurate picture of journal impact. The few revealed preference studies published in the field of Economics & Business (cf. Tahai & Meyer, 1999; Dubois & Reeb, 2000; Baumgartner & Pieter, 2003) focused on a very limited group of journals. This paper therefore presents a revealed preference study for more than 800 journals in the broad field of Economics & Business. It also introduces a new data source (Google Scholar) and a new citation metric (Hirsch's h-index, Hirsch 2005) to accommodate the critique levelled at ISI's JIFs and provides a benchmarking exercise of the two data sources and metrics.

DATA SOURCE AND METRICS

ISI Web of Knowledge versus Google Scholar

Seglen (1997) and Cameron (2005) provide good overviews of the problems with the ISI Web of Knowledge as a data source. These problems mainly revolve around ISI's limited coverage, especially in the Social Sciences and Humanities. Previous studies have highlighted issues such as: the lack of coverage of citations in books, conference and working papers as well as citations in journals not included in ISI; the lack of inclusion of journals in languages other than English in the ISI database; and the US bias in the journals included in the database (Harzing & van der Wal, 2008; Kousha & Thelwall, 2007/2008; Sanderson, 2008). Overall, there is considerable agreement that Google Scholar is a worthwhile alternative source of citation data, in particular in the Social and Information Sciences. Disadvantages of Google Scholar are its inclusion of non-scholarly citations, double counting of citations, less frequent updating, uneven coverage across disciplines and less comprehensive coverage of older publications/citations (Harzing & van der Wal, 2008).

The problem of non-scholarly citations and double counting is found to be fairly limited and attenuated by the use of robust citation metrics such as the h-index (Meho & Yang, 2007; Harzing & van der Wal, 2008; Vaughan and Shaw, 2008). The last three limitations are not relevant for this paper as we focus on a discipline that has good Google Scholar coverage and on citations to papers between 2000 and 2005.

Citations Metrics Used

Several commonly mentioned problems (for a summary see Seglen, 1997 and Cameron, 2005) with the ISI JIF are the use of a 2-year citation window, which for many disciplines is too short (Leydesdorff, 2008) and various technical issues related to the calculation of the JIF. First, whilst the denominator in the JIF (the number of

articles published) only includes normal articles (so called “source” items), the numerator includes citations to all publications in the journal in question, including editorials, letters, and book reviews (Cameron, 2005). This means that citations in these latter publications are “free” as the increase in the numerator is not matched by an increase in the denominator. Second, the JIF calculates the mean number of citations to an article in the journal in question. However, many authors have found that citation distributions are extremely skewed (e.g., Seglen, 1997). Individual highly cited papers can have a very strong influence on the mean JIF.

In this paper we use a relatively new citation metric: the h-index. The h-index was introduced by Hirsch (2005:1) and is defined as follows: “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.” Therefore, the h-index provides a combination of both quantity (number of papers) and quality (impact, or citations to these papers) (Glänzel, 2006). The h-index has resulted in a flurry of articles in journals such as *Scientometrics* and *Journal of the American Society for Information Science and Technology*, including articles proposing further refinements (cf. Bornmann, Mutz & Daniel, 2008) and has generally received a positive reception. Examples of the application of the h-index to journals are still scarce. Only brief notes have been published (cf. Braun, Glänzel & Schubert (2005) and Saad (2006)) and no study has covered more than a limited set of journals or provided a systematic comparison between different data sources and metrics (though see Saad’s (2007) working paper for a comparison between the eigenfactor (www.eigenfactor.org) and h-index).

The h-index has several advantages over the ISI JIF. First, it does not have a fixed time horizon.

The metrics used in the present paper were computed in October 2007 over a five-year period (2001-2005). However, any time horizon could be used, rather than focusing on citations in one particular year to the two preceding years as is the case with the ISI JIF. Second, the h-index attenuates the impact of one highly-cited article, because the h-index is not based on mean scores. Therefore, analogous to its use for authors, the h-index for journals provides a robust measure of *sustained* and *durable* performance of *journals*, rather than articles. Third, a journal that publishes a larger number of papers has a higher likelihood of generating a higher h-index, since every article presents another chance for citations. This is a disadvantage when evaluating the standing of individual *articles* in a journal (or an individual academic based on this metric) as this measure should not be dependent on the number of articles published in that journal. However, a journal that publishes a larger number of high-impact papers has a bigger impact on the field (see also Gissvold, 1999). Given that impact on the field is what we attempt to measure in this article, we argue this feature of the h-index is an advantage rather than a disadvantage.

METHODS

Data source

Since our aim was to cover a broader range of journals than in most previous studies we used Harzing’s Journal Quality List (Harzing, 2007). This list includes a collation of 20 different rankings of 838 journals in the broad area of Business and Economics. It appears to be quite influential: a search for the terms “Journal Quality List” AND Harzing results in more than 500 Google hits and the list has been cited more than 20 times in ISI listed journals (data for April 2008). The publisher of this list informed us that it is downloaded more than 10,000 times a year and that it draws interest from all over the world.

Procedures

The metrics used in this paper were calculated using Publish or Perish (<http://www.harzing.com/pop.htm>), a software programme that retrieves and analyses academic citations using Google Scholar as a data source. Searches were conducted in the first week of October 2007. We also searched for spelling variations (e.g., British vs. American spelling, the use of “and” vs. the use of “&”) and abbreviated journal titles (e.g., all SIAM journals). If a title included common words, e.g. *Journal of Management*, we conducted searches with the ISSN. As Google Scholar’s results for ISSN searches seem to be rather erratic, this alternative was only used if the ISSN search provided a comprehensive result. The results of all search queries were inspected for incomplete or inconsistent results. This left us with only two dozen journals (out of 838) that had substantially incomplete coverage and for which metrics could not be calculated. For other journals our inspection might have overlooked occasional missing articles, but this is unlikely influence robust measures such as the h-index unless they are highly cited. We have no reason to believe that this was the case.

Our Google Scholar searches included citations to articles published between 2001 and 2005. This timeframe was chosen to be comparable with a 5-year average for the JIFs of the last five available years (2003-2006). These JIFs refer to citations in articles published between 2001 and 2005. Supplementary analyses with regard to the extent of concentration of citations within a particular journal were conducted in October 2007 with the general search function of ISI, which allows the user to rank articles by citation.

RESULTS OF THE BENCHMARKING ANALYSIS

Overall comparison of JIF and h-index

There are 536 journals in the Journal Quality List that have both an ISI JIF for 2003-2006 and a Google Scholar (GS) h-index. The Spearman correlation between the ISI JIF and the h-index – used because both the JIF and h-

index have non-normal distributions - is strong and very significant: 0.718 ($p < 0.000$). Given that these two sets of indices have different data sources (ISI Thomson JCR versus Google Scholar) and provide different metrics (a mean citations-per-paper count over 2 years for the ISI JIF and a combined quantity/quality measure over 5 years for the h-index) this strong correlation is quite remarkable.

The Journal Quality List includes journals in 15 different sub-disciplines. There are seven sub-disciplines that have a substantial number (more than 60) journals included in the JQL: Economics, Finance & Accounting, General Management & Strategy, Management Information Systems & Knowledge Management, Management Science & Operations Research/Management, Marketing, and Organization Studies/Behavior and Human Resource Management/Industrial Relations (HRM/IR). Taken together these seven sub-disciplines cover 75% of the journals in the Journal Quality List.

Table 2 provides statistics for these seven sub-disciplines and shows there is significant variability in terms of the proportion of ISI-indexed journals in the different fields, ranging from a low of 30-43% for Finance & Accounting, Marketing, and General Management & Strategy to a high of 74-80% for Economics, Management Information Systems/Knowledge Management, and Management Science & Operations Research/Management. The sub-disciplines also differ in terms of the strength of correlation between the h-index and the JIF, varying from 0.633 for Organization Behaviour/Studies & HRM/IR to 0.891 for General Management & Strategy, but in all cases this correlation was highly significant.

Table 2: Summary statistics

Sub-field	No. of journals in the JQL	No. of ISI-indexed journals	Spearman correlation b/w h-index & JIF
Economics	168	122 (74%)	0.732***
Finance & Accounting	94	28 (30%)	0.721***
General Management & Strategy	63	27 (43%)	0.891***
Mgmt Information Systems & Knowledge Mgmt	81	61 (75%)	0.774***
Mgmt Science; Operations Research/Mgmt	87	70 (80%)	0.733***
Marketing	65	25 (38%)	0.841***
Organization Behaviour/Studies; HRM & IR	71	45 (63%)	0.633***
Others	209	158 (76%)	0.764***
Total	838	536 (64%)	0.718***

*** $p < 0.000$

Many journals, especially in Finance & Accounting, Marketing, and General Management & Strategy are not ISI-indexed. How do these journals compare with journals that *are* ISI-indexed? As expected, journals that are ISI-indexed have a significantly higher h-index (23.5 versus 11.5; $t = 15.002$, $p < 0.000$). However, there are more than 50 journals that ranked in the top 50% (16 and above) in terms of h-index, but are not ISI-listed. These journals are present in all disciplines, but are more frequent in the sub-disciplines that have a low ISI coverage. However, the single most distinguishing shared characteristic of these journals seems to be that they are published from Europe (usually by Blackwell, Elsevier, or Emerald) and generally have a European editor and a large proportion of non-US academics on the editorial board. Overall, nearly three quarters of the non-ISI indexed journals with a high h-index are European journals.¹

Reasons for divergence between JIF and h-index

Although there is a very strong correlation between the ISI JIF and the GS h-index, there are some notable cases where the two diverge. The reason for this divergence falls into five general categories, which provide a clear illustration of the advantages and disadvantages of the respective metrics.

Field-specific differences in immediacy index distort the ISI JIF

Most of the major outliers with a high JIF in comparison to their h-index are Psychology journals. Similar to Science journals, they have very high immediacy index; i.e. citations to these journals occur quickly after publication. For example, the 2006 immediacy index for the *Annual Review of Psychology* (4.091) is more

¹ We do not wish to imply that the ISI selection process has a bias against European journals. European editors might display a self-selection bias and simply do not submit their journals for inclusion into ISI. Of course one reason for this could be the perceived bias against non-US journals. The impact factor of the *American Economic Review* is more than ten times as high as that of the *American Economic Review* (0.335). This means that when comparing these two journals over a 2-year period (as is done for ISI JIFs) the *Annual Review of Psychology* will always show a higher impact factor than the *American Economic Review*, whereas the difference will be much smaller if we consider a 5-year period (as is done for the GS h-index). Hence a clear advantage of the GS h-index (and any metric based on a longer time-frame) is a fairer comparison across disciplines.

Individual highly-cited papers distort the ISI JIF

A key disadvantage of the JIF is that individual highly-cited papers can distort the JIF in individual years and hence many journals with a concentrated citation pattern show high JIFs in comparison to their h-index. A good example is *SIAM Review*, which had an average JIF of 2.75 between 2001 and 2003 and a JIF of 2.67 in 2006. However, in 2004 and 2005 its JIF stood at 6.12 and 7.21 respectively, causing a very high average JIF between 2003 and 2006. The very high JIFs for 2004 and 2005 were nearly entirely caused by a very large number of citations to **one** particular journal article published in 2003 (“The structure and function of complex networks” by MEJ Newman). In October 2007 this particular article had been cited 998 times, twelve times more than the next highest cited article published in 2003. In fact, in October 2007 the Newman paper alone makes up for 80% of the citations to *SIAM review* in 2003; the other twenty papers published in 2003 together have only 249 citations. This example clearly shows the danger of relying on mean-value metrics, such as the JIF, as they can be heavily influenced by individual outliers.

Citation in materials not covered by ISI increases the GS h-index

Google Scholar has a much broader coverage than ISI, including books, conference and working papers as well as a wide range of journals not included in ISI. As a result journals that garner a large proportion of their citations from these sources will generally have a relatively high GS h-index in comparison to their ISI JIF. Examples are abundant in all disciplines. Articles in journals such as *The American Economic Review* and *Research Policy* are cited very heavily in working papers (e.g., papers from the *National Bureau of Economic Research* or the *Tinbergen Institute*) and government policy documents. Articles in the various *IEEE Transactions* and *Communications of the ACM* are often cited in conference proceedings, which are the most important publication outlets in this field, but are not included in the ISI citation count. The *Journal of Business Ethics*, *Human Relations* and *International Journal of Human Resources Management*, all published out of Europe, have a high number of citations in European journals not indexed in ISI. These examples clearly show that the GS h-index provides a more comprehensive picture of a journal’s impact beyond the relatively narrow scope of ISI listed journals.

Number of papers published limits the GS h-index

The h-index is influenced to some extent by the number of papers that a journal publishes. A journal that publishes a larger number of papers has a higher likelihood of generating a higher h-index since every article presents another chance for inclusion in the h-index. Hence journals that publish a limited number of papers will generally show a GS h-index that is low compared to their ISI JIF. One example is the *Journal of Economic Literature*, which publishes a relatively small number of articles per year (15-20), so that even though most of these are highly cited, it will be difficult for the journal to achieve a very high h-index.² This is almost an exact counter case to the *American Economic Review*, which publishes around 160-170 articles per year that on average are not as highly cited as articles in the *Journal of Economic Literature*.

Overall, however, the *American Economic Review* has a much larger **total** number of articles that are highly cited. We therefore argue that the h-index correctly identifies the journal’s more substantial contribution to the field of Economics. On the other hand, when evaluating individual academics based on articles published in these two journals, one should clearly take this difference into account, and would probably assign a higher importance to publication in JEL. In this case, different metrics clearly serve different purposes.

Minor idiosyncratic reasons for divergence

ISI’s rather idiosyncratic calculation of the JIF includes citations to non-source material in the numerator, but not in the denominator. Hence journals with lively editorial/letter/book review sections display an ISI JIF that is high in comparison to their GS h-index. For instance, more than half of the *Academy of Management Review* papers are classified as either editorials or book reviews. Normally, this would not result in a significant distortion of the JIF as non-source materials tend not to be highly cited in Management journals (in contrast to for instance journals such as *Science* and *Nature*). However, the paper-length introductions to the many special issues are also classified as editorials and these pieces tend to be highly cited.

² It should be noted that *Journal of Economic Literature* published 91 papers between 2001 and 2005. Therefore, the journal could have achieved an h-index close to that of the *American Economic Review* or could be in the top-3 journals in terms of h-index. Clearly, ISI’s calculation method has the potential to distort the JIF and we would argue the GS h-index provides a more accurate measurement of impact.

One of the more striking cases of a journal with a high ISI JIF in comparison to its GS h-index is *Human Resource Management*. Thomson’s search query for this journal’s JIF was revealed to include a substantial number of homographs referring to *Human Resource Management Review*, *Human Resource Management Journal* as well as books with *Human Resource Management* in their title. As a result the JIF for *Human Resource Management* had been erroneously inflated. At 0.64 the recently released 2007 JIF for *Human Resource Management* is very substantially lower than the 2002-2006 JIF average of 2.00. It is possible that equally generic journal titles might suffer from the same problem. This – admittedly idiosyncratic – example shows that a comparison of different sets of metrics can help to spot errors in either of them.

DISCUSSION AND CONCLUSIONS

We showed that there is substantial agreement between the ISI JIF and the GS h-index for most sub-disciplines. Therefore, for those sub-disciplines that have limited ISI coverage (Finance & Accounting, Marketing, and General Management & Strategy) the GS h-index could provide an excellent alternative for the 56-70% of journals not covered in ISI. However, even for other sub-disciplines the additional coverage provided by Google Scholar could be useful.

Where the ISI JIF and the GS h-index diverged this was generally caused by one of four factors. First, the 2-year time frame of the JIF artificially rewards journals with a high immediacy index. Second, the sensitivity of the JIF to individual highly cited papers artificially inflates the JIF in comparison to the h-index. Third, the broader coverage of Google Scholar caused h-indices to be higher than JIFs for journals that receive a large proportion of their citations from policy documents, working papers, books, conference proceedings – none of which are included in ISI – or from journals that are not ISI indexed. Finally, the h-index is influenced by the number of papers published and hence journals that publish a lot of papers have a better chance to reach a high h-index, reflecting their broader impact on the field. In sum, the GS h-index addresses some of the statistical limitations underlying the JIF, and is more suitable to measure a journal's

(the current 3rd ranked journal has a h-index of 80) if all of its papers were highly cited. In fact, JEL has a very respectable h-index **in spite of** its limited number of papers and is the 34th ranked journal and in the top-5% of journals by h-index. wider economic or social impact rather than its impact on an academic audience only. As such we argue that the GS h-index provides a more accurate and comprehensive measure of journal impact and at the very least should be considered as a supplement to ISI-based impact analyses.

However, even though an assessment of journal impact based on the journal's GS h-index might be more accurate and comprehensive than relying only on an ISI-based impact analysis, we express strong caution against a single-minded focus on journal impact in evaluating individual scholars' research output. Whilst journal impact can certainly be used as **one** of the criteria to evaluate research output, reducing the evaluation to **one single** number is unlikely to provide a complete picture of a scholar's real impact. Many studies have established that highly-cited articles get published in journals that are not considered top journals in the field, and a substantial proportion of the articles published in top journals fail to generate a high level of citations (cf. Starbuck, 2005, Oswald, 2007 and Singh, Haddad & Chow, 2007). Hence using journal proxies to evaluate the impact of individual articles can lead to substantial attribution errors.

A more fundamental question is whether citation by other academics is the **only** relevant measure of impact. Another factor that could be considered in applied areas of research is whether the research in question "makes a difference" by providing insights into fundamental managerial or societal questions.

However, this assessment might be quite difficult to make and will always include some element of subjectivity. True managerial or societal impact might also not be apparent in the short term. Hence, although individual article impact and broader managerial and societal impact should be included in the evaluation of research output wherever appropriate and possible, most universities will by necessity place some emphasis on the use of journal impact proxies. In this article, we provided a broader perspective on journal impact and hope this will lead to a more valid and equitable assessment of academics' research output.

REFERENCES

- ABS (2007). Journal Quality Guide, downloaded from <http://www.the-abs.org.uk/?id=257&PHPSESSID=37610dc334308500ccd06adce0bcf043>.
- Braun T, Glänzel W, & Schubert A (2005). A Hirsch-type index for journals. *The Scientist* 19, 8.
- Bornmann, L., Mutz, R. & Daniel, H.-P. (2008) Are there better indices for evaluation purposes than the *h* index? A comparison of nine different variants of the *h* index using data from biomedicine. *Journal of the American Society for Information Science and Technology*, 59(5), 830-837.
- Baumgartner, H. & Pieters, R. (2003). The structural influence of marketing journals, a citation analysis of the discipline and its subareas over time. *Journal of Marketing*. 67, 123-39.
- Cameron, B.D. (2005). Trends in the Usage of ISI Bibliometric Data, Uses, Abuses, and Implication. *Portal: Libraries and the Academy*. 5, 105-125.
- DuBois, F. L. & Reeb, D. (2000). Ranking the international business journals. *Journal of International Business Studies*. 31, 689-704.
- Gisvold, S.E. (1999). Citation analysis and journal impact factors – is the tail wagging the dog? *Acta Anaesthesiol Scand*. 43, 971-973.
- Glänzel, W. (2006). On the opportunities and limitations of the H-index. *Science Focus*. 1, 10-11.
- Harzing, A.W.K. (2007). *Journal Quality List*, 28th Edition, <http://www.harzing.com/>
- Harzing, A.W.K. & Wal, R. van der (2008). Google Scholar as a new source for citation analysis?, *Ethics in Science and Environmental Politics*, 8, 62-71.
- Hirsch, J.E. (2005). An index to quantify an individual's scientific research output, [arXiv:physics/0508025](https://arxiv.org/abs/physics/0508025) v5 29 Sep 2006.

- Kousha, K.; Thelwall, M. (2007) Google Scholar Citations and Google Web/URL Citations: A Multi-Discipline Exploratory Analysis, *Journal of the American Society for Information Science and Technology*, 58, 1055-1065.
- Kousha, K; Thelwall, M. (2008) Sources of Google Scholar citations outside the Science Citation Index: A comparison between four science disciplines. *Scientometrics*, 74(2), 273-294.
- Leydesdorff, L. (2008) *Caveats* for the use of citation indicators in research and journal evaluations. *Journal of the American Society for Information Science and Technology*, 59, 278-287.
- McDonald, S.; Kam. J. (2007). Ring a ring o' roses: Quality journals and gamesmanship in Management Studies. *Journal of Management Studies*, 44, 640-55.
- McGarty, C. (2000). The citation impact factor in social psychology: a bad statistic that encourages bad science. *Current Research in Social Psychology*, 5, 1-16.
- Meho LI, Yang K (2007) A New Era in Citation and Bibliometric Analyses: Web of Science, Scopus, and Google Scholar. *Journal of the American Society for Information Science and Technology*, 58, 1-21.
- Mingers, J. & Harzing, A.W.K. (2007). Ranking journals in Business and Management: A statistical analysis of the Harzing Dataset. *European Journal of Information Systems*, 16, 303-316.
- Oswald, A. J. (2007). An examination of the reliability of prestigious journals: Evidence and implications for decision-makers. *Economica*, 74, 21-31.
- Reedijk, J. (1998). Sense and nonsense of science citation analyses: comments on the monopoly position of ISI and citation inaccuracies. Risks of possible misuse and biased citation and impact data, *New J. Chem*, 767-770.
- Saad, G. (2006). Exploring the h-index at the author and journal levels using bibliometric data of productive consumer scholars and business-related journals respectively. *Scientometrics* 69, 117-120.
- Saad, G. (2007). Convergent Validity Between Metrics of Journal Prestige: The Eigenfactor, Article Influence, and H-index Scores. Submitted for publication.
- Sanderson, M. (2008) Revisiting h measured on UK LIS and IR academics, *Journal of the American Society for Information Science and Technology*, 59, 1184-1190.
- Seglen, P.O. (1997). Why the impact factor of journals should not be used for evaluating research. *British Medical Journal*, 314, 497-502.
- Singh, G., Haddad, K.M., & Chow, C.W. (2007). Are articles in "top" management journals necessarily of higher quality? *Journal of Management Inquiry*, 16(4), 319-331.
- Starbuck, W.H. (2005). How much better are the most-prestigious journals? The statistics of academic publication. *Organization Science*, 16, 180-200.
- Tahai, A. & Meyer, M. (1999). A revealed preference study of management journals' direct influences, *Strategic Management Journal*, 20, 279-296.
- Vaughan, L.; Shaw, D. (2008) A new look at evidence of scholarly citations in citation indexes and from web sources. *Scientometrics*, 74(2), 317-330.