Use and Misuse of Bibliometric Measures for Assessment of Academic Performance, Tenure and Publication Support

Abstract: Bibliometric methods such as journal impact factor and article influence score based on number of citations were developed to measure and compare the quality of journals listed in citation indexes. Yet, they are increasingly being used nowadays for research assessment, hiring, tenure and academic promotion, research funding and publication support even though such metrics have not been developed to measure the quality of individual researchers or scientific articles. In this paper, we review the use of journal impact factor, cited half-life, article influence score and h index for academic performance assessment, academic promotion and publication support by Turkish universities and the Turkish Scientific and Technological Research Center (TUBITAK). Examples are provided regarding the consequences of using bibliometric measures beyond what they were originally designed for, and some recommendations are offered.

Key words: research assessment, academic promotions, journal impact factor, article influence score, h index

Introduction

This paper addresses the misuse of bibliometric measures for research assessment, academic promotion and monetary support for academic publications. Although most of the examples in this study come from the Turkish higher education system, such use is quite widespread in other countries as well and thus merits further investigation. For instance, the criteria set for academic promotions by the Turkish Higher Education Council (HEC) are almost entirely based on the number of papers published in ISI-indexed journals and the number of citations thereto. The Turkish Scientific and Technological Research Center (TUBITAK) provides monetary support to the authors of such papers on the basis of impact factors of journals in which their papers appeared. Similarly, the Research Council of Thailand provides monetary support to researchers simply by multiplying the number of papers they authored with the impact factors of journals in which they appeared (Arendt, 2010). Therefore, it is useful to look at more carefully as to what bibliometric measures such as journal impact factor, article influence score and h index exactly measure and whether they are suitable to use as criteria for academic promotions and publication support.

It should be mentioned at the outset that bibliometric measures alone are not the sole criteria for research assessment in scientifically more developed countries. Rather, such assessment is primarily based on peer review. For example, in the United Kingdom (UK) an experts panel evaluates the quality of research output of universities/departments as part of the Research Excellence Framework (REF) and allocates research budgets accordingly. In fact, panel members are not even allowed to use journal impact factors or any journal ranking system in evaluating academic publications (Sgroi & Oswald, 2013, p. F257). Peer review is also used in Turkey but only after the candidates satisfy the requirements of the number of papers/citations in ISI-indexed journals, presumably because bibliometric measures seem more “objective” to HEC authorities than the outcome of peer reviews. In the full paper we will provide the details of how HEC and TUBITAK used JIF, article influence score and cited-half life to make decisions on tenure and the amount of money to be paid to the authors of papers that appeared in ISI-indexed journals. The examples in this extended abstract are limited to journal impact factors and h index.

Use of Citation Indexes in Research Evaluation

Thomson Reuters’s citation indexes are primarily used for academic performance evaluation in many countries including Turkey. Yet, using the products of a commercial company in research evaluation
turned out to some problems, at least in Turkey. Up until about 10 years ago, the then ISI has been the sole publisher of citation indexes. When Elsevier’s Scopus entered the market with more journals indexed, Thomson Reuters decided in 2006 to increase its number of journals through what is called “regional expansion”. This decision was welcomed by many countries having a few journals in Thomson Reuters’s citation indexes. Turkey was no exception, as it increased its number of ISI-indexed journals from 5 to about 80! Needless to say, this increase cannot simply be explained by an unusual surge in the scientific level of Turkey. Yet, papers that appeared in those journals helped many academicians to satisfy the tenure requirements, not to mention the monetary support offered by TUBITAK to the authors of the same papers.

What exactly, then, does journal impact factor (JIF) measure? JIF measures how many citations an “average paper” gets in a certain period. As the San Francisco Declaration on Research Assessment (DORA) clearly states, JIF has been developed to help librarians select journals and it cannot be used to evaluate the quality of a paper (San Francisco, 2012). The disadvantages of using JIF in research evaluation are well documented in the literature (see, for instance, Casadevall & Fang, 2014; Marx & Bornmann, 2013; Seglen, 1997). Among them are: (a) citation distributions are skewed; (b) JIF varies by disciplines and can be manipulated by editorial policies; and (c) the data used to calculate JIF are not transparent and contributions other than research articles are also considered when calculating JIF. Of the 11,500 journals listed in Thomson Reuters’s citation indexes, 43% have JIFs between 0 and 1 (Al & Soydai, 2014). It is likely that considerable number of those low impact journals might have been added after the decision of regional expansion. Thomson Reuters, too, is against the use of JIF to measure paper quality (Marx & Bornmann, 2013, pp. 62-63). Therefore, taking the average JIF says almost nothing about the quality of an individual paper, let alone predict how many, if any, citations it would get in the coming years.

Use of H Index in Research Evaluation

Proposed by Hirsch (2005) as an alternative to the more traditional JIF, h index is meant to say something about the life-time scientific success of a researcher by means of productivity and impact. A researcher has an h index of n if s/he published n papers in ISI-indexed journals each of which receiving at least n citations. H index became very popular in a short period of time as it was easy to calculate and has since been used to measure the performance of not only researchers but also of universities, publishers or even single articles (Schubert, 2009). In time, it has also been used for academic performance evaluation.

However, h index also has some shortcomings. As h index is based on citation data and the citation rates of papers vary from discipline to discipline, the h indexes of researchers also vary. For example, the top most cited paper in Science published between 2008 and 2012 garnered over 1,000 citations whereas its equivalent in Economics received only 60 citations (Sgroi & Oswald, 2013, p. F256). In fact, the average JIFs tend to vary even in the subfields of Science. There is a considerable difference between average JIFs of Chemistry journals as opposed to that of Mathematics, which is due to the fact that the number of researchers (hence the potential number of researchers who would cite a given paper) and the number of journals in which they can publish are unequal. The average h indexes of researchers in these disciplines clearly reflect this. Moreover, h index is closely related with time: while senior professors continue to increase their h indexes based on citations they receive for their older publications even if they may not necessarily have current publications, junior ones who carry out research in relatively small disciplines and in social sciences and humanities need more time to not only publish more papers but also garner citations thereto in order to boost their h indexes. As h indexes are not normalized according to seniority and research fields, this limits the use of h indexes to measure the performance of researchers and compare them across academic positions (assistant, associate and full professors) and across disciplines. In addition, h index does not take co-authors into account (Hirsch, 2007, p. 19193). Some researchers think that h index does not meet some logical requirements and is not a first rate intellectual achievement but, rather, a “clever find”
Conclusions

Just because bibliometric measures such as JIFs and h indexes are readily available through Web of Science, Scopus or Google Scholar do not make them ideal measures to use for research evaluation and academic performance. JIFs, especially, seems to become “the poor man’s citation analysis” (Marx & Bornmann, 2013). The “fatal attraction” of bibliometric measures (Van Raan, 2005) that were not developed for research evaluation, tenure and publication support may have adverse effects on academic careers of researchers (Hudson & Laband, 2013, p. F201). We have already referred above to the San Francisco Declaration on Research Assessment (DORA) advising not to use bibliometric measures for research assessment. More recently, the Board of Directors of IEEE, “the world’s largest professional association for the advancement of technology” (ieee.org), adopted the statement that concludes: “. . . bibliometric performance indicators should be applied only as a collective group (and not individually), and in conjunction with peer review following a clearly stated code of conduct” (IEEE, 2013, original emphasis). We should pay heed to such recommendations and take the combination of peer review and bibliometrics as the “ideal way of research evaluation” (Bornmann & Leydesdorff, in press).

References

Hirsch, J.E. (2007). Does the h index have predictive power? PNAS, 104(49), 19193-19198.