Evaluating journal quality: A review of journal citation indicators and ranking in business and management

John Mingers

16 February 2017

Highlights

•Six quantitative indicators of journal quality are reviewed theoretically for their biases.
•They are then studied empirically on a sample of business and management journals.
•The results confirm the theoretical conclusions.
•A comparison with the ABS journal ranking list is made.
•No one indicator is considered superior to all the others.

Abstract

Evaluating the quality of academic journals is becoming increasing important within the context of research performance evaluation. Traditionally, journals have been ranked by peer review lists such as that of the Association of Business Schools (UK) or though their journal impact factor (JIF). However, several new indicators have been developed, such as the h-index, SJR, SNIP and the Eigenfactor which take into account different factors and therefore have their own particular biases. In this paper we evaluate these metrics both theoretically and also through an empirical study of a large set of business and management journals. We show that even though the indicators appear highly correlated in fact they lead to large differences in journal rankings. We contextualise our results in terms of the UK's large scale research assessment exercise (the RAE/REF) and particularly the ABS journal ranking list. We conclude that no one indicator is superior but that the h-index (which includes the productivity of a journal) and SNIP (which aims to normalise for field effects) may be the most effective at the moment.

Introduction

The evaluation of research performance, whether at the level of individuals, departments, research groups or whole universities, is becoming ever more important and the results of exercises such as the UK's Research Excellence Framework (REF) have major consequences in terms of funding and individual academics' careers. The primary driver of an evaluation is an assessment of the quality of an individual research output, generally a journal paper. The evaluation can be done by peer

review, as in the REF, or citations can be used as a proxy for quality – although they are really indicators of impact. The focus on quality of research has led to a focus on the quality of the publishing journal itself. There are several reasons for this: helping researchers decide where to target their papers; competition between the journals; and in many cases illicitly using the quality of the journal as a proxy for the quality of the papers published in it.

Journal quality, in turn, can also be evaluated either by peer review or by citation indicators such as the journal impact factor (JIF). Peer review has been the primary form in the past for journal ranking lists such as that of the Association of Business Schools (ABS) (Association of Business Schools, 2010).1 Many of these lists for business and management are available from the Harzing website (2009). Some of these lists, such as ABS, are a hybrid in that they use citation indicators to inform the peer review.

The practice of judging a paper by the journal in which it is published has become endemic within large scale evaluations, such as the UK's REF, where huge numbers of papers need to be graded but, as we shall see, this is not a practice to be recommended. Within business and management, in preparation for the 2014 REF, the ABS Guide was used by Schools to choose both papers and individual academics to be submitted, despite extensive criticism of the Guide from UK academics (Hussain, 2013, Mingers and Willmott, 2013, Walker et al., 2015, Willmott, 2011). It should be noted that the Business and Management REF Panel has repeatedly stated that they do not use any journal lists, and they have informally issued some data justifying this position, but this has not stopped the wholesale use of lists within business schools. This paper will discuss the results primarily within the UK context, but these large scale research evaluations also occur in Australia (Northcott & Linacre, 2010), New Zealand (Hicks, 2012) and Italy (Rebora & Turri, 2013).

These developments increase the importance of journal quality indicators, whether used in combination with peer review (as in the ABS list) or used instead of peer review. It is vital that the indicators available are accurate, robust, transparent and unbiased so that users, especially non-bibliometricians, can use them confidently (Wouters et al., 2015).

For many years the journal impact factor (JIF) was the predominant journal metric despite considerable criticism, but recently there has been a spate of new ones including the Eigenfactor, the h-index, SJR and SNIP and it is important to understand how these differ from each other,

and the degree of their validity (Moed, 2015, Straub and Anderson, 2010).
The purpose of this paper is to evaluate the indicators that are currently available in terms of these four criteria – accuracy, robustness, transparency and unbiasedness. It is evident that any metric has its own particular biases; that is, it will tend to favour certain kinds at the expense of others – that is after all the point of measuring something. Some of these biases will be explicit and indeed designed-in. Others will be implicit, perhaps not recognised, and may be undesirable. The review will analyse the theory of the different indicators, looking for their explicit and implicit biases, and then test these observations on a sample of journals from the business and management area. There are four sections which cover: a review of the different indicators; methodology and data collection; empirical results and comparisons with the ABS list; and recommendations.

Section snippets

Review of journal citation indicators

The use of citations to track the performance of journals was initiated by Garfield (1955) and he established the first citation indexes (Science Citation Index)2 and the company, the Institute for Scientific Information (ISI). Although, prior to that, the first analysis of papers citing a journal's publications occurred in 1927 (Gross & Gross, 1927) and Shepard's Citations is a legal citing

Methodology and data

We wished to compare the various indicators empirically on a sample of business and management journals and then compare the results with the ABS journal ranking. One of the problems is that the indicators are not all available from the same source – the JIF and Eigenfactor come from the WoS; the h-index from the Scimago website; and the SJR and SNIP from Scopus. Clearly this introduces problem of consistency as the databases do not cover the same set of journals and therefore have differences

Principal components

Given the interesting pattern of correlations, it is useful to conduct a principal components analysis to look at the relationships between the variables. Fig. 2 is a plot of the first two component loadings. PC1 Does not discriminate well between the indicators although those normalised for number of papers have higher values. PC2 distinguishes clearly

between these types of indicators with those un-normalised having positive values. We have not included the immediacy index as this is
Comparing journal indicators with peer review journal lists
In practice, at the moment, most journal ranking is actually done through peer reviewed lists such as the ABS list, or the Australian Business Dean's Council (ABDC) list14 which itself is a development of the more extensive Excellence in Research for Australia (ERA) list15 (Hall, 2011), although these may include some use of bibliometric indicators in their compilation.
Conclusions
This paper has considered the main journal impact indicators that are currently available through citation databases as these are the primary ones that are used in practice for decisions about journal ranking lists, destinations for research papers, jobs, promotions, and submissions to research evaluation programmes.
There are several general issues to be noted in terms of the appropriate use of these metrics. First, citation data is always highly skewed and this calls into question the validity
Acknowledgment
We would like to thank all the reviewers for their very helpful comments.
References (117)
S. Alonso et al.
h-Index: A review focused in its variants, computation and standardization for different scientific fields
Journal of Informetrics
(2009)
L. Bornmann et al.
The use of percentiles and percentile rank classes in the analysis of bibliometric data: Opportunities and limits
Journal of Informetrics
(2013)
L. Bornmann et al.
Methods for the generation of normalized citation impact scores in bibliometrics: Which method best reflects the judgement of experts
Journal of Informetrics
(2015)
R. Costas et al.
The h-index: Advantages, limitations and its relation with other bibliometric indicators at the micro level
Journal of Informetrics

(2007)
F. Franceschini et al.
Analysis of the Hirsch index's operational properties
European Journal of Operational Research
(2010)
B. González-Pereira et al.
A new approach to the metric of journals' scientific prestige: The SJR indicator
Journal of Informetrics
(2010)
V.P. Guerrero-Bote et al.
A further step forward in measuring journals' scientific prestige: The SJR2 indicator
Journal of Informetrics
(2012)
D. Hicks
Performance-based university research funding systems
Research Policy
(2012)
MaN. et al.
Bringing pagerank to the citation analysis
Information Processing and Management
(2008)
J. Mingers
Problems with SNIP
Journal of Informetrics
(2014)