



Publisher interest in citations: An analysis of Scopus database

N G Satish

Professor, Institute of Public Enterprise, Osmania University Campus, Hyderabad - 500 007, Email: satish.ng@ipeindia.org

Received: 01 May 2021; revised: 26 August 2021; accepted: 04 September

Publisher interest in citation to journal articles is an important variable that affect scholarly communication. The focus of citation studies is often on individual journal titles and Journal Impact Factors. The present study analyses the citation related data in the context of publishers represented in Scopus for the years 2016-2018. The results indicate a significant skewness in citations per document for 25 major publishers. Science academies and society publications, along with those publishers with fewer journals represented in Scopus, seem to be on the losing end in harnessing citation impact. Analysis points to the possible influence of publisher branding in the context of citation impact.

Keywords: Journal metrics; Publisher metrics; Citation analysis

Introduction

Citations to documents published in journals are considered an indicator of its quality. There have been, however, several contestations to this argument¹. Journal titles distinguish themselves on citations, referred to as Journal Impact Factor (JIF). JIF initiated by Garfield is a measure of average number of citations to articles published in a journal during the previous two years². Studies have shown that not all citable documents in journals get cited for various reasons³. In the recent years citation indices have increased the number of years considered for such calculations, and Scopus has also introduced a new prestige-based ranking called SCImago Journal Rank (SJR)⁴.

Though citation as a measure of individual worth has come under criticism, it is in the self-interest of the publishers to maintain and encourage the concept. Citation to articles provides brand building opportunity to publishers and journals⁵. Citation index, among other things, has become central to this activity.

Against this background an analysis was carried out for the Scopus indexed journals with the following objectives:

- To understand the publisher-wise distribution of journals in Scopus;
- To analyze publisher-wise distribution of total citations and citations per document;
- To examine difference among the publishers in terms of citation yield; and,

- To discuss possible strategies adopted by the publishers to sustain the citation advantage.

Methodology

Publisher and citation related information was collated from the Scopus indexed list of journals for the year 2018 available in Scopus (extracted in early 2020) through <https://www.scimagojr.com>. Publishing houses sometimes operate in different countries through their subsidiaries (e.g. Sage Publications India Pvt Ltd., Elsevier Urban and Partner sp) and in such cases they are pooled under one heading. In some cases publishers form joint ventures (e.g. Wiley-Liss Inc, Blackwell-Wiss.-Verl, Brunner - Routledge (US), Oxford & IBH Publishing) and such joint ventures, for the purpose of the current analysis, were normalized by grouping the titles under the more prominent publisher name such as Wiley, Blackwell, Routledge, Oxford, etc.

The corresponding data collected in the context include citable documents published by the journals, total citations and average citations per documents accrued to them during 2016-2018. The analysis used ANOVA to determine whether there is any statistically significant difference between the means of independent groups. Fisher's LSD was used for post-hoc analysis.

Fisher's LSD method creates confidence intervals for all pairwise differences between factor level means while controlling the individual error rate to a specified significance level. Fisher's LSD method then

uses the individual error rate and number of comparisons to calculate the simultaneous confidence level for all confidence intervals. This simultaneous confidence level is the probability that all confidence intervals contain the true difference. Data were analysed using SPSS software.

Results and analysis

Scopus claims an independent review mechanism for journal selection for indexing in its database. New suggestions from the publishers are said to be evaluated by the international experts using broadly defined quantitative and qualitative measures. Apart from the minimum criteria of availability of ISSN, English language abstract, and publicly available publication ethics in the journal, the criteria mentioned include - journal policy, along with type of peer review; geographical distribution of editors; authors; content; Journal standing in terms of citedness of journal articles in Scopus; editor standing; publishing regularity; and online availability are said to be considered for evaluation. Usage of abstract and full text is also considered for retention of the journals in the list once they are selected⁶. Thus, visibility, commercial prospect, and acceptable production standards are the criteria that stand out for inclusion of journals in Scopus.

Journal titles included in Scopus are published by academies, societies, local and international commercial publishers. Table 1 presents details on

clustering of journals across publishers and associated details. In 2018, Scopus indexed 24,690 journals. Publisher stakes in Scopus, in terms of titles included, range from single journal to as many as 2,114 of Elsevier imprint. Single title from the publisher makes up almost one-fourth (5,877 titles, 23.88%) of the journals included in Scopus. Those with 1-4 journals make up 33.33% (6, 855 titles) of the total journals.

The data shows that the top three publishers own 4,769 (19.37%) of the total journals in Scopus. These are Elsevier, Taylor & Francis, and Springer. Each of these business houses published over 1,000 plus indexed journals. The next three publishers, in descending order of titles indexed – Sage, Wiley, and Blackwell - owned Journal titles ranging from 500-999 making up 9.14% of the journals indexed in Scopus. These six top publishers cumulatively publish 36.04% of the total citable documents and have accounted for 42.29% of the citations accrued during 2016-2018 period.

A total of 19 distinct publishers owned journals ranging from 100-499 indexed in Scopus. These cumulatively work out 16.54% of the total journals in the database. Thus, the top 25 publishers managed about 45% of the total journals in Scopus. Together they published 50.99% of the total citable documents and accrued slightly over 60% of the total citations during 2016-2018 period.

At the other end of the journal distribution in Scopus are those publishing one journal. There are

Table 1 — Clustering of Journals and corresponding Citable Documents (2016-2018), and Citations (2016-2018)

Journals in Scopus	No Publishers	Titles	% Titles	Citable_Docs	%citable Docs (3 Yrs)	Total_Citations	%Total Citations (3 Yrs)
1000+	3	4769	19.37	1731397	26.88	5630226	31.30
500_999	3	2253	9.14	590335	9.16	1977578	10.99
400	1	466	1.89	140448	2.18	320420	1.78
300	3	982	3.99	175077	2.72	539274	3.00
200	6	1467	5.96	309342	4.80	779503	4.33
100	9	1153	4.70	338596	5.26	1684877	9.37
50_99	18	1184	4.81	451651	7.01	1723711	9.58
40	6	264	1.07	189324	2.94	856754	4.76
30	21	691	2.81	123413	1.92	283894	1.58
20	37	846	3.44	202866	3.15	545574	3.03
10	85	1139	4.63	320249	4.97	800779	4.45
5_9	190	1193	4.85	342316	5.31	947320	5.27
4	105	420	1.71	130355	2.02	199103	1.11
3	218	654	2.66	154919	2.40	297775	1.66
2	625	1250	5.08	267300	4.15	466993	2.60
1	5879	5877	23.89	974233	15.13	936342	5.20
	7209	24608	100.00	6441821	100.00	17990123	100.00

Source: Computed by the author

5,877 of those making up 23.89% of the total. They contributed 15.13% of the articles to Scopus and got a citation yield of 5.20% of the total.

Publishers with 1 to 4 journals indexed in Scopus make up 6,827 (33.33%) in number and contributed 23.69% of citable articles. They had accrued during the years only 10.57% of the total citations.

The above trends indicate that some commercial publishers with a large chunk of journals indexed in Scopus accrue more citations as opposed to those at the other end where the citable documents and the citations accrued do not match correspondingly. Varying number of titles owned by publishers, their corresponding citable documents in Scopus, and citation data is presented in Table 1. The distribution is explicit in being skewed towards a few top publishers both in terms of journal ownership, total citable documents, and total citations.

Major publishers represented in Scopus are shown in Table 2. As could be seen, Elsevier has 8.59% of the total journals, 17.86% of citable publications, and

24.92% of the citations. There seems to be an undue advantage for the Scopus publishers in coverage and citations.

ANOVA was computed to understand whether the citations per document for the journals differ significantly across the publishers. Citation per document was computed by dividing total citations by citable documents. As some of the journals in the list were new additions, the corresponding data was not available, and they were eliminated from the analysis. Initially the sample was divided into two groups, namely major publishers with 100 plus journals in Scopus and the rest of the publishers. The ANOVA summary is presented in Table 3.

The analysis reveals a significant F Ratio ($F=421.975$ (1,24444), $MSE=22.064$ $p<.000$) indicating that the mean citations per document accrued to major publishers with large number of journals in Scopus tend to be significantly more citations per document compared with the minor publishers. Publishing houses through their

Table 2 — Major publishers, their journals (2018), and the corresponding citable documents and citations in (2016-2018)

Publisher	Titles	% Titles	Citable_Docs	%citable Docs (3 Yrs)	Total_Citations	%Total Citations (3 Yrs)
Elsevier	2114	8.59	1150426	17.86	4483735	24.92
Taylor & Francis	1271	5.16	186786	2.90	334671	1.86
Springer	1384	5.62	394185	6.12	811820	4.52
Wiley	890	3.61	315644	4.9	1224471	6.81
SAGE	818	3.32	127099	1.97	284025	1.58
Blackwell	545	2.21	147592	2.29	469082	2.61
Kluwer Academic	466	1.89	140448	2.18	320420	1.78
Cambridge Univ. Press	344	1.40	48743	0.76	85548	0.48
Routledge	323	1.31	31384	0.49	44689	0.25
Oxford Univ. Press	315	1.28	94950	1.47	409037	2.27
Emerald	288	1.17	32046	0.50	57779	0.32
Lippincott Williams & Wilkins	261	1.06	100035	1.55	291623	1.62
Bio Med Central	250	1.02	86988	1.35	290011	1.61
Inderscience	230	0.93	17006	0.26	13778	0.08
Walter de Gruyter	223	0.91	25771	0.40	29833	0.17
Hindawi	215	0.87	47496	0.74	96479	0.54
Bentham Science	170	0.69	18945	0.29	32052	0.18
IEEE	149	0.61	99895	1.55	520881	2.90
Brill Academic Publishers	147	0.60	8145	0.13	3846	0.02
W. B. Saunders	122	0.50	37805	0.59	112512	0.63
Haworth Press	118	0.48	9609	0.15	10478	0.06
Carfax Publishing	115	0.47	16363	0.25	31201	0.17
Medknow Publications	115	0.47	26747	0.42	30771	0.17
Maney Publishing	114	0.46	12428	0.19	14967	0.08
Nature Publishing Group	103	0.42	108659	1.69	928169	5.16
Major Publishers	11090	45.05	3285195	51.00	10931878	60.79
Others	13518	54.95	3156626	49.00	7058245	39.21
Total	24608	100.00	6441821	100.00	17990123	100.00

Source: Computed by the author

Table 3 — ANOVA Summary: Publisher Groups vs Citations Per Document

	No of Journals	Mean citations per document		Sum of Squares	df	Mean Square	F	Sig.
Major publishers with 100+ journals	11191	2.3149	Between Groups	9310.608	1	9310.608	421.975	.000
Others	13255	1.0762	Within Groups	539341.017	24444	22.064		
Total	24446	1.6432	Total	548651.625	24445			

promotional policies seems to have an influence of how the journal content is perceived by the users.

The analysis of citations per document was carried forward taking top 25 publishers who play a major role in Scopus contents to understand whether significant variations manifest among these publishers within the group. ANOVA carried out for this group (Table 4) also reveal a significant F Ratio ($F=20.261$ (24,11166) $MSE=40.069$ $p<.000$) indicating that some publishers within this elite group tend to accrue significantly higher citations per document than the others. Post hoc analysis of the ANOVA using Fisher's LSD was carried out to understand which of the publishers differ significantly within the group. In the current context Fisher's LSD test is a set of individual publisher-wise t tests with all the others to know the statistical difference in distribution of citations per document. The test computes the pooled SD from all the groups. The resultant mean difference could be – and + in the matrix. Significant '–' value indicates that the citations per document in journals of that publisher is significantly less than the other in contention. The LSD analysis results are presented in Table 5

As could be seen from the results, six publishers stand out among the 25 top contributors to Scopus. These are BioMed Central (250 titles indexed in Scopus), Blackwell (545), Elsevier (2114), IEEE (149), Nature Publishing (103), and Wiley (890). These publishers register a significantly higher citations per document compared to the others in this select group. BioMed Central, IEEE and Nature Publishing are specialized narrowly focused publishers with relatively lesser number of journals among the select top contributors. Though Blackwell journals tend to score significantly higher citations per document than the 10 other publishers in this group, they also accrue significantly less than the five others in the group, including Elsevier. Wiley journals score significantly higher on 15 other publishers in this group. Wiley's citations per document in Scopus are significantly less than Elsevier. Only Elsevier tends to score consistently high on all the others, excepting the

Table 4 — Distribution of titles and ANOVA summary: Major Publishers vs Citation per Document

Publisher	No of Titles	Mean Citations Per Document
Bentham Science	169	1.07
Bio Med Central	274	3.08
Blackwell	557	2.56
Brill	147	0.34
CUP	343	1.31
Carfax	119	1.68
Elsevier	2107	3.55
Emerald Group	286	1.57
Haworth	149	1.03
Hindawi	215	1.70
IEEE	173	5.01
Inderscience	230	0.69
Kluwer	476	2.12
Lippincott	261	2.16
Maney Publishing	114	0.87
Medknow	115	1.13
Nature	102	12.00
OUP	315	2.54
Routledge	323	1.27
Sage	817	1.94
Springer	1283	1.87
Taylor & Francis	1268	1.53
W. B. Saunders	122	2.51
Walter De Gruyter	339	0.83
Wiley	887	2.86
Total	11191	2.31

five among them. Two of these are Nature and IEEE, both specialized publishers. And, for the other three the difference is not statistically significant. This indicates relative greater influence exercised by Elsevier in Scopus citation world. Needless to mention that Scopus is an Elsevier product.

Publisher interest in citations does not seem to play out differently in Science Citation Index (SCI) Expanded, another Citation Indexing product. SCI Expanded indexed 9,500 journals at the time of this study. These bear imprint of 1,752 distinct publishers (<https://mjl.clarivate.com/collection-list-downloads>). As was computed from the data, the top 20 of these overlap with that of Scopus in proportion of publisher

Table 5—Post hoc LSD test for the publishers
(* P<.05, **<.01, Blank squares indicate Non Significant relations)

BENTHAM SCIENCE	2.00*	1.49*	2.00	-1.49	-2.00	-1.49	-2.48*	1.50*	2.05*	1.38*	-3.93*	-10.93*	-1.47*	-1.79*
BIOMED CENTRAL	1.49*	2.22*	1.77*	1.40*	-2.48*	1.50*	2.05*	1.38*	-3.93*	-1.93*	2.39*	1.81*	1.20*	2.24*
BLACKWELL	2.73*	2.22*	1.77*	1.40*	-2.48*	1.50*	2.05*	1.38*	-3.93*	-2.44*	1.8*	1.30*	.693*	1.73*
BHILL	2.22*	1.77*	1.40*	-2.48*	1.50*	2.05*	1.38*	-3.93*	-2.44*	-1.35*	-4.66*	-11.66*	-1.53*	-2.52*
CUP	2.73*	2.22*	1.77*	1.40*	-2.48*	1.50*	2.05*	1.38*	-3.93*	-3.70*	-3.33*	-10.69*	-1.19*	-1.55*
CARFAX	1.49*	2.22*	1.77*	1.40*	-2.48*	1.50*	2.05*	1.38*	-3.93*	-1.87*	-1.46*	1.00*	1.67*	2.012*
ELSEVIER	2.48*	1.98*	1.85*	1.87*	-1.98*	1.98*	2.52*	1.85*	-3.44*	-1.46*	2.86*	2.28*	1.61*	2.71*
EMERALD GROUP	1.5*	1.54*	1.35*	1.35*	-1.98*	1.98*	2.52*	1.85*	-3.44*	-1.46*	2.86*	2.28*	1.61*	2.71*
HAWORTH	2.05*	1.54*	1.35*	1.35*	-2.52*	1.98*	2.52*	1.85*	-3.98*	-3.44*	1.01	-10.43*	1.67*	2.71*
HINDAWI	1.35*	1.35*	1.35*	1.35*	-1.85*	1.85*	3.98*	3.31*	-3.31*	1.01	4.32*	-10.97*	1.52*	-1.29*
IEEE	3.94*	1.93*	2.44*	4.66*	1.46*	3.33*	3.70*	3.33*	-3.31*	1.01	4.32*	-10.30*	3.069*	-1.17*
INDERSCIENCE	2.39*	1.88*	1.77*	1.77*	-2.86*	1.46*	3.70*	3.33*	-4.32*	2.89*	2.85*	-6.99*	3.138*	2.15*
KLUWER	1.77*	1.77*	1.77*	1.77*	-1.42	1.42	1.39*	1.13	-4.32*	2.89*	2.85*	-11.31*	3.47*	2.18*
LIPPINCOTT	1.81*	1.81*	1.81*	1.81*	-1.39*	1.39*	1.13	1.13	-2.89*	1.43*	1.47*	-9.88*	2.50*	-1.82*
MANEY Publishing	2.20*	1.69*	1.69*	1.69*	-2.67*	1.39*	1.13	1.13	-2.89*	1.43*	1.47*	-9.84*	3.47*	1.29*
MEDKNOW	1.95*	1.44*	1.44*	1.44*	-2.42*	1.43*	1.13	1.13	-2.85*	1.47*	1.47*	-9.84*	1.32*	1.32*
NATURE	10.93*	8.92*	9.44*	11.66*	8.45*	10.32*	10.30*	10.30*	-4.13*	1.31*	9.88*	-11.13*	-1.63*	-1.99*
OUP	1.47*	2.20*	1.24*	1.24*	-1.00*	1.00*	1.52*	1.52*	-3.88*	6.99*	11.31*	-10.87*	-1.41*	-1.73*
ROUTLEDGE	1.81*	1.30*	1.30*	1.30*	-2.28*	1.30*	1.52*	1.52*	-3.88*	6.99*	11.31*	-10.87*	10.13*	9.13*
SAGE	1.14*	1.59*	1.59*	1.59*	-1.61*	1.61*	1.52*	1.52*	-3.74*	3.74*	10.06*	10.73*	10.47*	9.49*
SPRINGER	1.20*	1.53*	1.53*	1.53*	-1.67*	1.67*	1.52*	1.52*	-3.069*	1.25*	11.12*	10.06*	10.13*	11.16*
TAYLOR & FRANCIS	1.54*	1.03*	1.03*	1.03*	-2.01*	1.03*	1.52*	1.52*	-3.13*	1.19*	11.12*	10.06*	10.13*	11.16*
W. B. SAUNDERS	2.14*	2.14*	2.14*	2.14*	-2.01*	2.14*	1.52*	1.52*	-3.47*	1.82*	1.67*	1.24	1.01*	1.70*
WALTER DE GRUYTER	2.24*	1.73*	1.73*	1.73*	-2.71*	1.73*	1.52*	1.52*	-2.50*	1.82*	1.63*	1.24	1.01*	1.70*
WILEY	1.79*	2.52*	1.55*	1.55*	-69*	1.29*	1.83*	1.16*	-4.17*	2.18*	1.99*	1.60*	1.33*	2.02*

representation. These include Elsevier (16.20% of the total), Springer (13.24%), Wiley (9.81%). These three make up almost 40% of the coverage. The other publishers with more than 100 titles in the source are – Taylor and Francis, Sage, BioMed Central, Lippincott, Williams and Wilkins, Oxford University Press, IEEE and Cambridge University Press. These top 10 publishers in this source publish 58.36% of the total journals indexed.

M/s Clarivate claims to be publisher-independent citation database, most probably referring to them (the index publisher) not having any interest in the journals indexed. However, given the skewed distribution of title ownership in the index the publishers-wise skewed citation distribution may hold good there as well.

Journal publishing is an expensive endeavour. In that competitive market, citation plays an important

promotional role. And hence the publisher interest in citations. Despite varying motivations to cite and several objections in the scholarly literature about its validity as quality measure, citations' utility to the publishers also seem to keep the practice in vogue. Some publishers actively participate in the process. In public accounts, M/s Elsevier have described the value they add to publications through their investment, including "coordinating the review, consideration, added text and references, and production and distribution mechanisms"^{7, 8}.

Most of the major publishers instruct authors on sharing and promoting their articles as an important part of research⁹. Along with fostering the exchange of scientific information they also seem to serve the publisher interest in promoting their journals. Commercial publishers exploit this feature well. This must be seen in the context of journal production which is an expensive commercial proposition with an eye on the profit margins. For some of these publishers, profit margin is estimated to be 30% and could be as high as 50% in some cases⁸.

There are at least two commercial services for the purpose of promotion and citation seeking currently in operations - Kudos and Impactstory. Kudos (<https://info.growkudos.com/>) aims to help expand readership of research publications and increase citations, via a structured process that includes writing a lay summary and using social media effectively. Impact Story (<https://profiles.impactstory.org>) facilitates creation of online profiles of research outputs to track the altmetric impacts. Citation has willy-nilly transformed into a marketing and brand building tool.

Conclusion

Citation index is an important discovery tool. By obtaining citation advantage through clustering of

titles by major publishers, and subsequently positioning their products as the high impact conveyors they influence both the scholarly information generation and their use. This phenomenon needs critical examination.

Acknowledgements

This study is a part of the grant-in-aid provided by NSTMIS Division of DST, New Delhi.

References

- 1 Knorr-Cetina K, The manufacture of Knowledge: An essay on the constructivist and contextual nature of science, (Oxford, Pergamon Press), 1981, 189 p.
- 2 Garfield E, The History and Meaning of the Journal Impact Factor, *JAMA*, 295(1) (2006) 90-93. <http://garfield.library.upenn.edu/papers/jamajif2006.pdf> (Accessed on 02 May 2021)
- 3 Callaway E, Beat it, impact factor! Publishing elite turns against controversial metric: Senior staff at leading journals want to end inappropriate use of the measure, *Nature* 535 (7611) (2016) 210–211
- 4 Description of Scimago Journal Rank Indicator, Available at <https://www.scimagojr.com/SCImagoJournalRank.pdf> (Accessed on 02 May. 2021).
- 5 Marland A, Journal publishing and marketing in an age of digital media, open access and impact factors, *Canadian Journal of Political Science*, 50(1) 2017, 77-95.
- 6 [Elsevier] Content Policy and Selection. https://www.elsevier.com/solutions/scopus/how-scopus-works/content/content-policy-and-selection_ (Accessed on 26 August 2021).
- 7 Poynder R, Elsevier's Alicia Wise on the RWA, the West Wing, and Universal Access, February 2012. Available at https://richardpoynder.co.uk/Wise_Interview.pdf (Accessed on 26 August 2021).
- 8 The STM Report: An overview of scientific and scholarly publishing - 1968-2018, (International Association of Scientific, Technical, and Medical Publishers), 2018. https://www.stm-assoc.org/2018_10_04_STM_Report_2018.pdf (Accessed on 26 August 2021).
- 9 [Elsevier] Sharing and promoting your article. <https://www.elsevier.com/authors/submit-your-paper/sharing-and-promoting-your-article> (Accessed on 26 August 2021).