

Note on Publishing Intensity and Publishing Breadth

John Sum, National Chung Hsing University, Taiwan
Chang-Han Song, National Chung Hsing University, Taiwan

The Asian Conference on Psychology & the Behavioral Sciences 2014
Official Conference Proceedings 2014
0350

Abstract

Given a set of leading scholars and a period of time, publishing intensity (PI) of a journal is defined as the total number of publications appeared in the journal that are co-authored by the leading scholars. Publishing breadth (PB) is defined as the total number of leading scholars who have contributed publications in that journal. While PI and PB have been applied in journal ranking, their dependency has not been investigated. In this regard, this paper presents empirical analyses on the relation between PI and PB, with focus on six areas namely Artificial Intelligence, Information Science and Library Science, Management, Anthropology, Geography, and Nursing. For each area, we first extract the list of journals from the JCR 2012 Version. Second, the list of leading scholars is compiled based on three rules: (1) a leading scholar must currently be an editorial member of a journal in which the journal is in our journal list, (2) a leading scholar must be affiliated with one of the Top 25 US universities compiled by US News, and (3) a leading scholar must have publications in the area during 1999 to 2003. The last rule ensures that a leading scholar has been active in the area for more than ten years. Third, based on the lists of journals and leading scholars obtained, we count from the Thomas Reuter WoK Database the PI and PB for each journal. Finally, we analyze the log-log relation between the PI and the PB of the journals in the list. Results show that log PI and the log PB have a linear correlation. The same result appears in all six areas. As the six areas have quite diverse natures, we argue that this log-linear relation is a common behavior across other research areas.

Introduction

Scholars who are planning to publish their research paper to journals may face a question: which journal is better? The measurement of the quality of a journal can be in many ways. All of these measurement schemes have a same goal: giving journal a rank position to let people know which one is better. Nowadays, there are various journal ranking approaches that are being used commonly. They may rank journal in many aspects, for example, "quality", "influence", "popularity", "reputation" or some other factors. However, these journal ranking approaches work rely on the indices it contained. A new journal ranking approach that can reflect the real quality is a must. In the pass decades, a considerable amount of literatures has been published on methods of journal ranking. However, every ranking method has its drawbacks. Journal ranking using opinion-based indices may be subjective. On the other hand, ranking using citation-based reflects the frequency with which the journal's articles are cited in the scientific literature. Nevertheless, the best way to characterize the productivity and prestige of journals is through the combination of various measures and indicators simultaneously (Bollen et al., 2006).

To solve the bias of the existing ranking approach, a better alternative would be the combination of different indices. Publication Power Approach mentioned two ranking indices, which are publishing intensity and publishing breadth. While PI and PB have been applied in journal ranking, their dependency has not been investigated. In this regard, this article has three purposes: to provide an empirical analyses on the relation between PI and PB, with focus on six areas namely Artificial Intelligence, Information Science and Library Science, Management, Anthropology, Geography, and Nursing.

There are various journal ranking approach that are commonly used nowadays, each journal ranking method can be composed of different indices. Take TR Impact Factor for example, it uses the number of citations to be its measure indices. For it is based on the citation number, we call it Citation-based index. It measures the quality and impact of a journal by its paper citation numbers.

In the paragraph below, we list the name of the ranking approach and the reference paper of the approach.

Citation-based index

- TR Impact factor ("Genetics Citation Index." Garfield E, Shepard IH, 1963), it uses JCR to be its database.
- H-index. Google Scholar Metrics uses the concept of h-index to rank journal as well.
- Some are opinion-based. ("An index to quantify an individual's scientific research output", Hirsch 2005)
- C-index ("A Bibliometric Index Based on the Collaboration Distance between Cited and Citing Authors.", Domingo-Ferrera, Maria Bras-Amorósa Josep, and Vicenç Torrab., 2011)
- G-index ("Theory and practise of the g-index", Leo Egghe 2006)
- Eigenfactor (Bergstrom, C. T., 2007)
- SCImago Journal Rank (SJR) (Guerrero Bote, Vicente P., Borja González-Pereira, and Félix de Moya-Anegón. ArXiv (2009)).

Opinion-Based Index

- Expert survey
- Author affiliation index (“IS Journal Quality Assessment Using the Author Affiliation Index”, Ferratt, Thomas W.; Gorman, Michael F.; Kanet, John J.; and Salisbury, Wm. David, 2007)

Behavior-Based Index

Holsapple provides a different perspective when doing journal ranking. The factors he used are publishing intensity, publishing breadth and publishing mode. (“A Publication Power Approach for Identifying Premier Information Systems Journals”, CW Holsapple, 2008)

Publication Power Approach (PPA) was proposed by Holsapple (Holsapple, 2008). It collects the publishing behaviors (includes how many papers a scholar published and in which journals in a given period) of a set of active scholars. Some people may say that it is a behavior-based approach, but in fact, there're so many cannot be separated or defined. So in here, we put it in the category of scholar opinion approach. The method is as follows:

Let M be the total number of active scholars, N be the total number of journals, I_j and B_j respectively be the publishing intensity and the publishing breadth of the j -th journal. The publishing intensity and breadth are defined as follows:

$$I_j = \sum_{i=1}^M a_{ij},$$
$$B_j = \sum_{i=1}^M \text{sign}(a_{ij}), \quad \forall j = 1 \dots N.$$

With the definition of publishing intensity and publishing breadth, CW Holsapple proposed a measure called publication power which is defined as follows (Holsapple, 2008) :

Publication power = publishing intensity * publishing breadth

Let publication power be U , publishing intensity be

$$U = I_j * B_j$$

The benchmark faculty examined here is composed of all full-time, tenured professors in IS at these schools in June 2006. These results in 73 benchmark faculty members are tabulated from June 2006 back to 1980. In the results reported here, only those journals with publishing intensity of at least 10 are included.

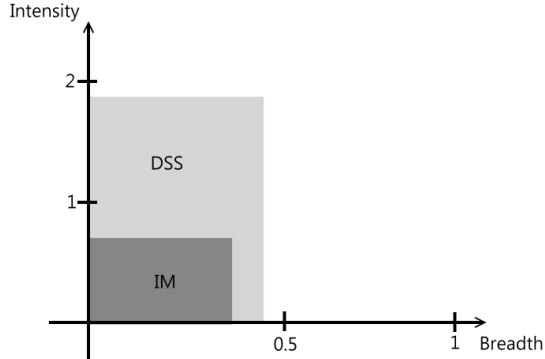
Example :

	Breadth		Intensity		Power	
Journal	Total(score)	Rank	Total(score)	Rank	Product	Rank
DSS	34(.47)	1	136(1.86)	1	4624	1

I&M	27(.37)	2	50(.68)	2	1350	2
-----	---------	---	---------	---	------	---

*A *breadth score* for a journal is obtained by dividing its publishing breadth by the total number of benchmark faculty members.

* A *intensity score* for a journal is obtained by dividing its publishing intensity by the total number of benchmark faculty members.



In this example, we can see that the publication power of DSS is higher than IM (4624>1350), so the ranking for DSS is No.1 and for IM is No.2

Hypothesis

While the author was collecting data for PPA, he used AAAI to be its benchmark scholars. We are curious that what the result will be if we change the benchmark scholars, for what is the standard to determine an association to be benchmark scholars is still debatable and there isn't always a prestigious association for every journal field. As a result, we choose IEEE CIS fellow to be our benchmark scholars and see whether the result may be different from the original one.

The definition of Publishing Intensity and Publishing Breadth

Let M be the total number of active scholars, N be the total number of journals, I_j and B_j respectively be the publishing intensity and the publishing breadth of the j-th journal. The publishing intensity and breadth are defined as follows:

$$I_j = \sum_{i=1}^M a_{ij}$$

Publishing intensity is defined as the total numbers of articles authored by the active scholars

$$B_j = \sum_{i=1}^M \text{sign}(a_{ij}), \quad \forall j = 1 \dots N.$$

Publishing breadth is defined as the total numbers of active scholars who have authored articles in the journal.

Here is an example showing how to calculate the publishing breadth and publishing intensity. For instance, there are two active scholars Prof. Ho and Prof. Sum respectively. There are the journals in the area of Technology Management, namely Journal of Information Systems (IS), Journal of Technology Management (TM) and Journal of Electronic Commerce (EC).

	IS	TM	EC
Prof. Ho	3 (articles)	1 (articles)	0 (articles)
Prof. Sum	2 (articles)	3 (articles)	2 (articles)
Breadth	2	2	1
Intensity	5	4	2

Publishing breadth is the total numbers of active scholars who have authored articles in the journal. For IS and TM, both scholars have published papers in these journals. So, their publishing breadth is two. For the journal EC, only Prof. Sum has published papers on it. So, its publishing breadth is one. As the publishing intensity is defined as the total numbers of articles authored by the active scholars, their values are clearly five, four and two respectively.

Exam based on Publication Power Approach

Publication power approach is a new ranking approach that has our interests. Publication power approach uses the product of two indices to be its final ranking index. However, no one explore the relation of publishing intensity and publishing breadth. Therefore, we follow the guide from the original paper to obtain our own statistic. The results are in the table 2. Interestingly, when we applied log model to both of the indices and then plot them, it came out to a graph like fig 1-6. As a result, we make a hypothesis: Publishing intensity and publishing breadth are log-linear-related.

Methodology

Rules for data collecting process

Talking about the data collecting process, we need to clarify our standard for filtering valid journals and active scholars first. For the valid journal, the quality of journal is not easy to measure, thus, we make a simple rule: only to include the journals that have been published for no less than 15 years. We believe that the longer time a journal has been published, the more scholars and readers it may attracted. On the other hand, it need lots of effort and research output to become a journal editor. However, to make sure that each editor has the higher reputation and more research,

we add two more rules about the working place and publishing time. In sum, the active scholars must fulfill these three rules: A journal editor who also works in top 25 universities and has publishing record between 1999 and 2003. The purpose of having publishing record from 1999 to 2003 is to exclude the scholars that only publish papers in recent years.

Steps for data collection

In this section we illustrate the adjusted method in different fields. The principles to perform the analysis are as follows:

Step 1: Find Journal List

For the benchmark journals, the first step is retrieve the list of journals from JCR database in 2012 social science edition in Tomson-Reuters Web of Knowledge (WoK). All the list of journals we used was extracted from WoK.

Step2: Find Qualified Journals

To make sure the quality of the journal, we exclude the journals that have published for less than 15 years. The reason why we do this is journals with older published years may attract more professors to publish their papers in it and thus attract more reader. After the filtering, the remaining journals become our qualified journals.

Step 3: Find Editors

In order to find the set of active scholars, we must find a way to choose scholars who make great effort to the research area. Thus, we decided to use editorial board to be the potential active scholars list. Those scholars may come from all over the world and would have made significant and huge contribution to the research field. As a result, choosing this set will make the result more reliable and more convincing. We retrieve the editorial name from the office website of each journal; it includes editors, associate editors, editor member/board and advisory editors.

Step 4: Find Qualified Editors (Active Scholars)

We search the editorial teams/board of those journals and check if those editors are in the top 25 public schools proposed by USNews. It is now known as the leader for ranking colleges, graduate schools and hospitals. The schools are UCLA, UC Berkeley, UC Davis, UCSD, UCSB, UC Irvine, Georgia, Michigan, Maryland, Wisconsin, Texas, Texas A&M, Florida, William and Mary, Penn State, Rutgers, Illinois, Washington, Virginia, North Carolina, Georgia Institute of Technology, Ohio State, Pittsburgh, Connecticut, Purdue, Clemson and Minnesota. The professors teaching in these schools have higher chance to produce high quality papers. By comparing the editorial list with the top 25 public schools list, if they are matched, they become out active scholars. Moreover, we check whether they've been published in this field for more than 10 years to make sure that those active scholars are not publishing their articles only in recent years.

Step 5: Data Collection Process

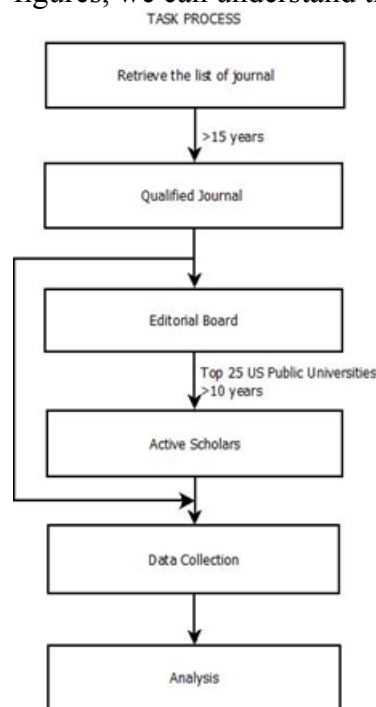
We use TR Web of Knowledge database in the data collection step. We match the list of active scholars and qualified journal to get the Publishing Intensity and Publishing Breadth.

Take Management field for example, 172 journals were found in the TR WoK database. 146 out of 172 journals published more than 15 years. By searching all the

editors in these 146 journals and matching them with top 25 public schools list, there still remained 643 editors. We traced the 643 editors' publishing history, came out that only 194 senior professors have published papers in Management field for more than 10 years (during 1999-2003). In other word, we compared about twenty eight thousands (194*146) items for management field. In the whole progress, this is the most time-consumed step.

Step 6: Analysis

Analyzing the data by Microsoft Excel and IBM SPSS software, we can calculate the linear regression equation and find the Slope, T-ratio and R-Squared value. In order to get the T-ratio, we set the confidence interval to be 95%. With these statistics and figures, we can understand the relation between LPI and LPB.



Data Analysis

After we collect the data, we enter the data analysis step. Linear regression is a statistic approach used to model the dependence of a scalar variable and one(or more) explanatory variables. In our case, there is only one explanatory variable, so it is called simple linear regression. We can represent it mathematically

$$\text{Intensity} = b \cdot \text{Breadth}^a$$

$$\log \text{Intensity} = a \cdot \log \text{Breadth} + c$$

R-Squared Value:

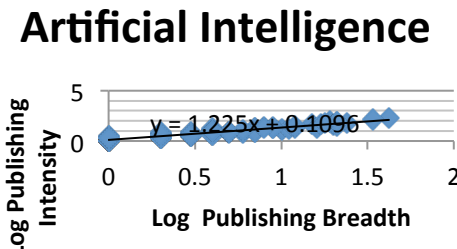
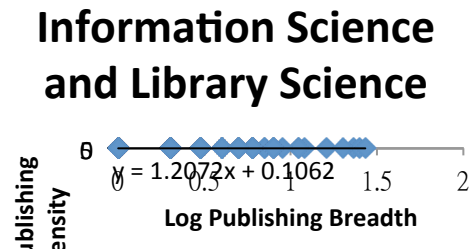
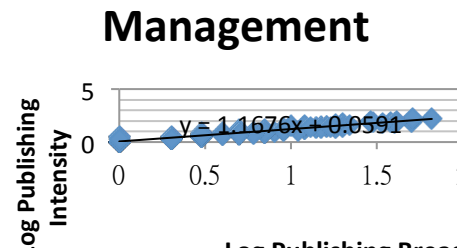
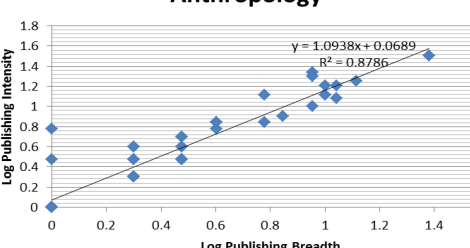
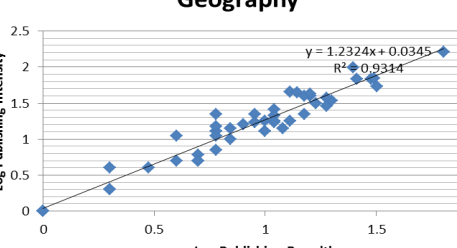
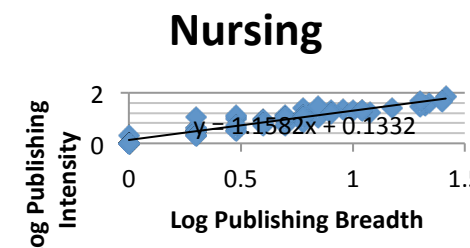
R-Squared is mostly being used for measuring the strength of correlation in linear regression model. Its value indicates how well the resulting line matches the original data point. From the statistic point of view, if R-Squared value of a data set equals to 1 means that the regression line perfectly fits the data. In other words, R-Squared value of the regression is relatively high indicates the points will be very close to the regression line. In the case of Management field, the R-squared values for the line is 0.946, suggesting that LPI and LPB are highly correlated.

T-Ratio:

T-ratio value is the indicator to determine the significance of regression coefficient. The significance of a regression coefficient is determined by dividing the estimated coefficient over the standard deviation of this estimate in a regression model. We can look for the appropriate $\alpha/2$ significance level to find the exact critical value from the t-distribution table. To find the significance of their relationships, we expect the t-ratio value to be greater than 2.

Result and Conclusion

Table1. The plot of log publishing intensity and log publishing breadth

 <p>Artificial Intelligence</p> <p>Log Publishing Intensity</p> <p>Log Publishing Breadth</p> <p>$y = 1.225x + 0.1096$</p>	 <p>Information Science and Library Science</p> <p>Publishing Intensity</p> <p>Log Publishing Breadth</p> <p>$y = 1.2072x + 0.1062$</p>
<p>Fig1 Log relations in the field of Artificial Intelligence</p>	<p>Fig2 Log relations in the field of IS &LS</p>
 <p>Management</p> <p>Log Publishing Intensity</p> <p>Log Publishing Breadth</p> <p>$y = 1.1676x + 0.0591$</p>	 <p>Anthropology</p> <p>Log Publishing Intensity</p> <p>Log Publishing Breadth</p> <p>$y = 1.0938x + 0.0689$ $R^2 = 0.8786$</p>
<p>Fig3 Log relations in the field of Management.</p>	<p>Fig4 Log relations in the field of Anthropology.</p>
 <p>Geography</p> <p>Log Publishing Intensity</p> <p>Log Publishing Breadth</p> <p>$y = 1.2324x + 0.0345$ $R^2 = 0.9314$</p>	 <p>Nursing</p> <p>Log Publishing Intensity</p> <p>Log Publishing Breadth</p> <p>$y = 1.1582x + 0.1332$</p>
<p>Fig5 Log relations in the field of Geography.</p>	<p>Fig6 Log relations in the field of Nursing.</p>

Journal field	Slope (t value)	b value (t value)	R-Squared Value
Artificial Intelligence	1.225 (25.943)	0.110 (3.076)	0.899
Information Science and	1.207 (21.193)	0.106 (2.612)	0.894

Library Science			
Management	1.168 (44.478)	0.059 (2.666)	0.946
Anthropology	1.094 (19.024)	0.069 (2.158)	0.879
Geography	1.232 (24.433)	0.035 (0.677)	0.931
Nursing.	0.881 (20.345)	0.133 (3.205)	0.881

Table2. The statistic of each field of journal, including slope, b value and R-square value.

Journal field	Journals	Qualified Journals	Active Scholars
Artificial Intelligence	115	90	208
Information Science and Library Science	85	73	97
Management	174	115	194
Anthropology	83	73	164
Geography	72	47	120
Nursing.	104	81	168

Table3. The statistic of the total journal number, valid journal number, total active scholar number.

This paper examined the two factors in public power approach, which are publishing intensity and publishing breadth respectively. The result showed that the publishing intensity and publishing breadth have linear relations between them. From our analysis, we can consider PI & PB are identical in 6 different journal fields. When doing the ranking approach, we can choose either PI or PB to be the ranking index. This indicated that different indices may have relations between them and it provided us more options to the combination of journal ranking indices.

Reference

- Bergstrom, C. T., West, J. D., & Wiseman, M. A. (2008). The EigenfactorTM metrics. *The Journal of Neuroscience*, 28(45), 11433–11434.
- Dalkey, N., & Helmer, O. (1963). An experimental application of the Delphi method to the use of experts. *Management science*, 9(3), 458–467.
- Donohue, J. M., & Fox, J. B. (2000). A multi-method evaluation of journals in the decision and management sciences by US academics. *Omega*, 28(1), 17–36.
- Hawkins, R. G., Ritter, L. S., & Walter, I. (1973). What economists think of their journals. *The Journal of Political Economy*, 81(4), 1017–1032.
- Holsapple, C. W. (2008). A publication power approach for identifying premier information systems journals. *Journal of the American Society for Information Science and Technology*, 59(2), 166–185.
- Rokach, L. (2012). Applying the publication power approach to artificial intelligence journals. *Journal of the American Society for Information Science and Technology*, 63(6), 1270–1277

