



This article has been submitted to Loughborough University's Institutional Repository by the  
author.

## BRIEF COMMUNICATION

# DOES BRADFORD-ZIPF APPLY TO BUSINESS AND MANAGEMENT JOURNALS IN THE 2001 RESEARCH ASSESSMENT EXERCISE?

Valerie Bence and Charles Oppenheim\*

Department of Information Science

Loughborough University

Loughborough

Leics LE11 3TU

- To whom correspondence should be addressed

## **BACKGROUND**

As part of research on journal submissions to the UK Research Assessment Exercise (RAE), data was compiled for a longitudinal analysis for Unit of Assessment (UoA) 43, Business and Management Studies (BMS). Data on the 1996 RAE submissions for this UoA was considered. For the 2001 RAE, academic Departments in the UK wishing to submit to the RAE were asked to nominate up to four items of published output per member of staff being submitted that represented the best work those individuals had published in the period. There was no obligation to just use journal articles, as all types of publication were acceptable, but in practice the majority of items returned (.....%) were journal articles. The remainder comprised reports, conference proceedings and similar outputs. Journal submission data obtained from the publicly available database (<http://www.hero.ac.uk>) following the 2001 exercise was checked, tabulated and analysed. This showed that for 2001, 7941 journal articles were submitted to 1489 journal titles<sup>1</sup>. This sample of journal titles could be said to represent the best UK BMS research for the period 1996-2001 (the period under consideration in the 2001 RAE). . The question arises whether journals in which articles for the RAE are submitted represent the “core” journals for that subject. One way of approaching this question is to subject the data to a Bradford-Zipf analysis. If the material shows a classic Bradford-Zipf distribution, it could be argued perhaps that the core journals in the field can indeed be identified using this method.

In order to address this issue and as part of the on-going research, the group of 1489 titles was used to test for a Bradford-Zipf distribution.

## **BIBLIOMETRIC LAWS**

The Bradford-Zipf distribution is one of the best known of the so-called bibliometric laws. These laws provide tools that can be used to answer many practical problems such as:-

Policy - How well are particular institutions doing in research?

Library planning - What is the "core" literature of a particular field?

Source assessment - What is the impact of a particular journal?

History of science - How did a particular field develop?

Sociology of science - Who is working with whom in the invisible college?

---

<sup>1</sup> No allowance has been made for articles submitted more than once, either within or between institutions. The results use a straight count only. This is mainly because if articles were submitted more than once, they were presumably assessed more than once, as part of each institution's submission. A separate analysis is being made of co-and multi-authored submissions.

**Bradford's Law** (or law of scattering) concerns the way these papers are distributed in journals (*Bradford, S.C. Documentation. Washington, DC, Public Affairs Press, 1948*). Bradford chose journals for his analysis because of their characteristics of the repetitive occurrence of themes and tendencies in specific fields of knowledge. If journals are ranked by the number of articles they contain on a given topic they can be divided into a central nucleus of the most important journals and a series of zones each containing the same number of articles as the nucleus (but each containing many more journals). It is used, for example, by librarians when deciding which journals to buy to best serve the research needs of their readers. They can estimate how many of the most productive journals they need to buy to yield a given fraction of items.

In addition, **Zipf's Law** (*Zipf, G. Human behaviour and the principle of least effort. New York, Hafner, 1972 573p.*) describes the frequency distribution of words in a given text, with familiar words being used many times and many words being used only once. Bradford's and Zipf's laws have been shown to be mathematically identical (Brookes, 1968) and so the distribution is often referred to as the Bradford-Zipf distribution (*Brookes, B. C. The derivation and application of the Bradford-Zipf distribution. Journal of Documentation 24, 1968, 247-265*).

Bradford's law is based on the fact that every scientific field is related, however remotely, to every other field, and when looking at the journals within a field there is always a small group of core journals that account for a substantial percentage (approx one-third) of the articles on that subject or discipline. Then, there is a second larger group of journals that account for another third, with a larger group accounting for the final third. (*Garfield, E. Bradford's law and related statistical patterns. Essays of an Information Scientist Vol 4 1979 p. 477*). However, it has been pointed out that "the Bradford-Zipf "law" does not allow one to predict the size of a core corpus of literature in a particular subject" (*Holmes, A and Oppenheim, C. 2001 Use of citation analysis to predict the outcome of the 2001 Research Assessment Exercise for Unit of Assessment (UoA) 61: Library and Information Management. Information Research, 6 (2)*)

Other bibliographic studies confirm that the dispersion of articles throughout a set of journal titles conforms to the distribution identified by Bradford. Most empirical investigations into Bradford's law come from within the natural and medical sciences, with few from the humanities or Social Sciences, however see (*Alypeku, W.O. Bradford distribution theory – compounding of Bradford periodical literatures in geography. Journal of Documentation, 33 1977, p. 296-304*) and (*Coleman, S. R. 1993 Bradford distributions of social science bibliographies varying in definitional homogeneity. Scientometrics, 27 (1) 75-91*).

This introduction of the concepts of 'core' and 'scatter' are fundamental to the study of knowledge production and boundary studies. Chubin observes that "if there was no 'scatter' scientists would be divided into small groups sharing the same interests, speaking only to each other, and reading and citing only each others work....both core and scatter are necessary, the former to permit scientific knowledge to cumulate and grow and the latter to prevent it from becoming ...a sect-like phenomenon." (*Chubin, D. E. The conceptualisation of scientific specialities. Sociological Quarterly 17 (3) 1976 p. 472 448-476*)

In addition to some of Bradford's early work ( ), Lawani, (*Lawani, S M. Bradford's law and the literature of agriculture. International Library Review, 1973, 5, 341-350*) applied the law to groups of articles in a graphic form and found that the line initially appears as an upward curve before it becomes linear. Journal title numbers are plotted logarithmically along the horizontal axis, and the cumulative sums of articles are plotted along the vertical axis.

However, Garfield points out that Zipf's law does not describe a Bradford distribution precisely, but a rough constant is yielded "only for the straight line portion of a Bradford curve". (*Garfield, E. Bradford's law and related statistical patterns. Essays of an Information Scientist Vol 4 1979 479*). The expectation is for the presence of a relatively small number of highly productive titles and a long tail of journal titles that contribute few articles.

## **DATA USED**

Journal titles submitted to the 2001 RAE for BMS were converted into raw data, which was tabulated into the columns required for Bradford-Zipf calculations as follows, column A - journal record number (1 being the most productive journal, 2 the second most productive and so on), column B – number of submissions to the given journal title, and column C - cumulative total of articles. Column A was then plotted against column C using semi-log paper using the approach adopted by Brookes. With such a volume of data, it is not possible to give it in full<sup>2</sup>, so Data is provided in the Appendix of the 100 most productive journals only.

## **RESULTS**

Figure 1 shows the Bradford-Zipf bibliograph for the data. For a classic Bradford-Zipf distribution, there should be a lengthy straight-line portion to the graph in the middle of the graph. No such straight line appears, and it therefore must be concluded that the Bradford-Zipf distribution does not apply to the journals submitted to the 2001 RAE in BMS. It is therefore difficult to identify the "core" journals in the field by this method.

---

<sup>2</sup> Copies of the data are available from the authors on request

## CONCLUSION

Two possible conclusions can be drawn from this result. The first is that the data set was not appropriate for a Bradford-Zipf analysis. Such analyses typically consider a comprehensive (or an attempted comprehensive) bibliography on a given topic. However, this collection of data represented just what was considered to be the best outputs (up to a maximum of four per person) by the Departments concerned, and so, by definition, the data is not comprehensive, and indeed makes no attempt to be comprehensive. The second is that –as has been argued by other authors (M.H. Hasso and C. Oppenheim, *Secondary services in Archaeology: an evaluation*, *J. Librarianship and Information Science of Iraq*, 1987, **12**, 1-18) – the Bradford-Zipf distribution is not a scientific law that is always obeyed, but rather a distribution that sometimes happens to follow a rule but sometimes does not, and this is another example of where it is not obeyed.