Can we rank scholarly book publishers? A bibliometric experiment with the field of history.
Zuccala, Alesia; Guns, R. ; Cornacchia, R.; Bod, R.

Published in:
American Society for Information Science and Technology. Journal

Publication date:
2015

Document version
Publisher's PDF, also known as Version of record

Citation for published version (APA):
This is a publisher ranking study based on a citation data grant from Elsevier, specifically, book titles cited in Scopus history journals (2007–2011) and matching metadata from WorldCat® (i.e., OCLC numbers, ISBN codes, publisher records, and library holding counts). Using both resources, we have created a unique relational database designed to compare citation counts to books with international library holdings or libcitations for scholarly book publishers. First, we construct a ranking of the top 500 publishers and explore descriptive statistics at the level of publisher type (university, commercial, other) and country of origin. We then identify the top 50 university presses and commercial houses based on total citations and mean citations per book (CPB). In a third analysis, we present a map of directed citation links between journals and book publishers. American and British presses/publishing houses tend to dominate the work of library collection managers and citing scholars; however, a number of specialist publishers from Europe are included. Distinct clusters from the directed citation map indicate a certain degree of regionalism and subject specialization, where some journals produced in languages other than English tend to cite books published by the same parent press. Bibliometric rankings convey only a small part of how the actual structure of the publishing field has evolved; hence, challenges lie ahead for developers of new citation indices for books and bibliometricians interested in measuring book and publisher impacts.

Introduction

Bibliometricians do not know very much about academic book publishers. Unlike journals, which have been studied intensively (Haustein, 2012), book publishers are much like fortresses: shadowy strongholds in the scholarly communication system. Scholars, particularly humanists, rely on them to print “the book” that will lead to tenure and career promotion (Cronin & La Barre, 2004; Dalton, 2006; Williams, Stevenson, Nicholas, Watkinson, & Rowlands, 2009). Universities are keen to promote their presses as additions to their scholarly reputation, and evaluators are growing curious about whether a publisher’s authority is equal to a proven measure of quality.

Publisher “quality” has previously been assessed using survey studies (Garand & Giles, 2011; Goodson, Dillman, & Hira, 1999; Lewis, 2000; Metz & Stemmer, 1996) and has also been related to library catalogue holdings (Donovan & Butler, 2007; SENSE-Research School for Socio-Economic and Natural Sciences of the Environment, 2009) and number of reviews per publisher and publisher reputation (Jordy, McGrath, & Rutledge, 1999). The concept of prestige, which is slightly different from “quality,” is often associated with academic publishing advice, specifically the selection of a university press by a scholar who has written a new book (Pasco, 2002; Pratt, 1993; Rowson, 1995).
Consideration is usually given to the age of the press and its original mission as well as the rank of its parent university (Gump, 2006). More often than not, ranking procedures for publishers are constructed on the basis of sales and/or yearly revenues (e.g., Publishers Weekly, 2013).

The focus of this research is to determine how publisher prestige may be indicated quantitatively using bibliographic citations to books from Scopus journal articles within the broader field of history. Knowledge in this regard can play a role in how books are selected for new and improved citation indices. Our study is made up of three parts. The first is a statistical exploration of highly cited publishers according to type and country of origin. The second part constructs a ranked list of 50 book publishers—university presses and commercial publishing houses—based on citation counts in journal articles versus mean citation counts per book. This is followed by a comparative analysis of mean citations to mean library holding counts, or “libcitations,” per book (White et al., 2009). Our statistics are derived from a newly constructed Scopus-WorldCat® relational database. Whereas the citation normally indicates the scholarly use of a particular book, a “ libcitation” is considered to be a measure of the book’s perceived cultural benefit (Torres-Salinas & Moed, 2009; White et al., 2009). Directed journal-to-publisher networks are presented in the final part of the study and recommended for revealing regional as well as specialty publishing patterns.

Background Literature

A Brief History of Journal Rankings

Bibliometric research has traditionally focused on journals and the ranking of journals based on their relative importance to a research field. Nisonger (1999) indicates that the first approach to ranking was undertaken by Gross and Gross (1927) for the field of chemistry, and, at the time of Nisonger’s own research (on library and information science journals), thousands of journal rankings had already been undertaken. Why do we rank? Initially it was to delineate a journal’s role in a formal communication network and assist librarians with collection-management decisions (Archambault & Larivière, 2009). Over time it has become an instrument for helping scholars with “manuscript submission decisions and with planning their publication agenda; for assessment of faculty performance regarding tenure, promotion, and annual raise decisions” (Nisonger, 1998, 1999, p. 1004). Garfield is often credited with translating these motivations into a full-fledged industry, with the Institute for Scientific Information (ISI) Journal Citation Index (now Thomson Reuters’ Web of Knowledge) and seminal work on ranking journals by frequency and impact of citations (Garfield, 1964, 1972).

In recent years, the bibliometrics community has been eager to improve our perception of journals with refinements to existing journal impact measures (Glänzel & Moed, 2002). Statistical measures generally lead to a ranking, but there can be calculation inaccuracies, which have been known to mislead (Moed & van Leeuwen, 1995, 1996). Moreover, a journal’s rank can change depending on the type of measure that is used. The first journal impact factor (IF), which is widely known, was introduced by Garfield (1964, 2006); whereas the source normalized impact factor (SNIP), developed at Leiden University (Moed, 2010) and the SCImago Journal Rank from Spain (González-Pereira, Guerrero-Bote, & Moya-Anegón, 2010) are both newer additions to the general indicator toolkit. All three indicators focus mainly on scientific journals that accumulate high citation counts. Certain journals published in the humanities are amenable to impact factors (see Elsevier, 2010), but many do not accumulate citation frequencies as we see in science and particular fields in the social sciences. Normalization techniques are imperative because a journal’s performance is not appropriately measured without considering the research field to which it belongs and the time to citation for papers cited in that field (Leydesdorff, Zhou, & Bornmann, 2013; Moed & van Leeuwen, 1996; Van Leeuwen, 2012).

Rarely have we seen a journal’s impact measure described in terms of prestige as opposed to rank; however, Ball (2006) gave this latter concept considerable attention in a Nature article. The article itself was based on a study carried out by Bollen and his colleagues (2006) at the Research Library of the Los Alamos National Laboratory in New Mexico. Although the journal impact factor may be seen as a measure of crude popularity, and the term crude holds more meaning nowadays with criticisms of the IF (e.g., Archambault & Larivière, 2009; Calver & Bradley, 2009); the Los Alamos team suggest that Google PageRank (PR) can be used as a complementary measure of prestige. The product of the two measures generates what is known as the Y-factor. This Y-factor can then be used to point out why “certain journals can have high IFs but low PRs (perhaps indicating a popular but less prestigious journal) and vice versa (for a high-quality but niche publication)” (Ball, 2006, p. 170). For instance, “among physics journals, the IF places Reviews of Modern Physics at the top of the list, but the Y-factor shifts the emphasis to rapid-publication journals. Physical Review Letters is the most influential, with a Y-factor of 5.91 × 10−2” (Ball, 2006, p. 170).

Haustein’s (2012) Multidimensional Journal Evaluation further emphasizes the value of employing and integrating complementary measures. To date, this book is one of the most comprehensive analyses pertaining to journals; referring to earlier research by Todorov and Glanzel (1988), Rousseau (2002), Van Leeuwen and Moed (2002), and Moed (2005a, 2005b) and detailing extensively why a single metric is not adequate. It is not within the scope of this paper to replicate Haustein’s (2012) work, but there is much to learn from her exploration. In sum, multiple factors can contribute to a journal’s rank, namely, its publication output, content, perception by readers and usage, citations received, and the management of the journal itself (i.e., the editorial board and publisher affiliation). This type of information is of equal relevance when we develop impact measures for book publishers.
From Journals to Books

Both a scientific journal and a scholarly book are often the product of the same “parent” press; thus (in some ways) book publishing is familiar but in other ways not. General impact measures for book publishers have simply not been developed because citation indices have always omitted full metadata for books. Garfield (1996), creator of the early ISI Journal Citation Index notes that for the social sciences and humanities “the failure to include monographs as sources may be a drawback in drawing conclusions about the impact of certain work.” It is almost as if book publishers have simply been forgotten, even though they “stand at a crucial crossroads in the production and distribution of knowledge in any society. They are in a position to decide what is ‘in’ and what is ‘out’ of the marketplace of ideas” (Coser, 1975, p. 14). Some countries or regions, such as Norway (Sivertsen, 2010) or Flanders, Belgium (Engels, Ossenblok, & Spruyt, 2012), have funding and evaluation systems that put emphasis on books and book publishers, but this practice has not yet been widely adopted. In Norway, for instance, the publisher is one of the factors used to determine a book’s weight, including whether the book should be included in an evaluation. The Norwegian system distinguishes between “level 1” and “level 2” publishers on the basis of “qualitative judgment and consensus among peers” (Sivertsen, 2010, p. 26).

Thomson Reuters’ approach to rectifying the book-evaluation problem has been to augment its Web of Knowledge with the addition of a new Book Citation Index (BKCI). Elsevier is doing the same with the addition of more than 75,000 book titles to Scopus (Adams & Testa, 2011; Elsevier, 2013; Thomson Reuters, 2013). Torres-Salinas, Robinson-García, Jime’nez-Contreras, and Delgado López-Cózar (2012) recently assessed the content of the Thomson Reuters’ BKCI to determine optimal indicators for a preliminary “Book Publisher’s Citation Report.” Thus far, the research team has identified a total of 19 indicators, with six focused specifically on ranking scholarly publishers: Three are related to productivity (i.e., total number of items indexed, number of books indexed, number of chapters indexed) and another three help to determine the publisher’s impact (i.e., total citations for all items, average citation per item, percentage of noncited items).

Exploratory work with the BKCI is underway (Leydesdorff & Felt, 2013; Torres-Salinas, Robinson-García, Campanario, & Delgado López-Cózar, 2013), yet what we often hear about publishers, especially university presses, is not that they require evaluation but that they are barely surviving as a result of financial cutbacks (Dalton, 2006; Greco, 2009; Thatcher, 1999). According to Greco (2009), “the vast majority of all university presses require subsidies from their home universities” and “very few end up in the ‘black’” (p. xi). This stands in sharp contrast to the journal industry, in which presses with active journal publications “are highly regarded and essentially ‘profitable’ (i.e., they generate a surplus)” (p. vii).

Despite this publishing crisis, humanities scholars from a variety of disciplines still feel pressure to publish a book. Cronin and La Barre (2004) found that, within research-intensive faculties for English and foreign language studies, “the book is still the principal coin of the realm” even though “equivalency can be established in a variety of ways” (p. 89). The expectation is that candidates for tenure should write at least one book; however, little evidence from the surveyed faculties (n = 101) pointed to clear guidelines for new faculty, nor was it clear that publishing a book is considered the best course of action. Cronin and La Barre (2004) suggest that a “change of some kind seems possible” given that new modes of electronic publishing are emerging (pp. 86–87).

For many scholars who do feel pressure to publish a book, selecting the best publisher is critical, if not as important as writing the book itself. Goodson et al. (1999) found that in tenure and promotion cases, in hiring decisions, in departmental gossip, and at APSA [American Political Science Association] meetings and other professional gatherings, with whom one publishes does matter. One colleague wrote us, “I vividly remember a conversation with a fellow author. . . . [who] refused to submit a manuscript to certain publishers because they were not prestigious enough. In my experience (about thirty years), there is a decided “unwritten” hierarchy of publishers, not only in political science, indeed, not only in academia.” Another noted, “My general sense is that the pecking order of publishers is quite clear within my subfield (at least between the best and the rest).” If there is a clear hierarchy among publishers, then which presses are viewed by members of the discipline as publishing the highest quality books? (p. 257)

Goodson et al. (1999) employed a list-based survey and asked scholars in the field of political science to rate the quality of known publishers. The five presses that received the highest mean quality ranking were Cambridge University Press, Princeton University Press, Oxford University Press, University of Chicago Press, and Yale University Press. When respondents were asked to evaluate only those presses/publishing houses with which they were familiar, Cambridge once again topped the list, followed by Oxford University Press and then Harvard University Press. Although the results were not surprising [note: Oxford and Harvard are generally thought to be “gold standards” in publishing (Dalton, 2006)], it is useful to mention that there are certain drawbacks to using survey questionnaires. Not only are they time consuming to administer, but it can be difficult to obtain readership participation or to construct a sample that fully represents all of a discipline’s subfields. The concept of “prestige” is also focused entirely on what scholars believe to be true about publishers, with no objective measure of which published books are actually used and cited most often in the research literature.

Garand and Giles (2011) later expanded on the survey approach by using an open-ended questionnaire (again with political scientists). In this study, a comparison was made
between scholars’ own choice of publisher (i.e., indicate the first, second, and third book publisher to which you would submit a manuscript) versus their readership preferences (i.e., identify book publishers whose books you read regularly or consider to be the best research in your area of expertise). Again, the results put Cambridge University Press and Princeton University Press at the top in terms of publishing choice and nothing changed with respect to readership: “Cambridge University Press and Princeton University Press [led] the way once more, with preference counts for Princeton University Press constituting slightly more than 80% of the preference counts for Cambridge University Press” (p. 377).

Improvements can always be made to surveys; however, it is perhaps time now to augment this type of research with bibliometric methods of evaluation. In fact, rankings based on surveys have been shown to relate positively to bibliometric measures, at least in research with journals (Rousseau, 2008). Both approaches have their merit: the first gives us a glimpse of “perceived” publisher reputation in light of faculty expectations; whereas the second quantitative approach, which allows for wider observation, is now becoming more feasible with recent developments to Scopus and Reuter’s new BKCI. Years of developing impact measures and ranking journals have prepared us for the task of ranking book publishers; we are (it is to be hoped) less inclined to employ misleading statistics and consider optimal ways of interlacing indicators.

The Publishing Industry

All insights gathered, regardless of analytic method, are significantly related to how the book publishing industry itself has evolved. Certain facets of the book publishing industry have been persistent, and Thomson (2005) refers to these as “the structure of the field” (p. 86). This basic structure comprises four criteria: (a) ownership status, (b) economic capital, (c) symbolic capital, and (d) geographical reach.

In terms of ownership status, there are clear differences between the university press and the commercial publishing firm. The university press is usually constituted as a department within a host university, and often it is registered as a charity or not-for-profit. Many university presses are also overseen by a committee: Cambridge University Press and Oxford University Press have, respectively, a Board of Syndics and a Board of Delegates (Thomson, 2005). Unlike the university press, the commercial publisher may be a private, largely family-owned business, or it could be a subsidiary of a larger corporation (e.g., as Longman is to Pearson). The commercial publisher is subject to financial objectives and commercial constraints. Moreover, it is not required to publish scholarly material and does not have any special tax status.

For both the university press and the commercial publisher, a high level of economic capital makes a difference in operations and production. At the commercial end, industry reports identify publishers that bring in millions of dollars of revenue annually. In Publishers Weekly (2013), Pearson from the United Kingdom is said to have retained its “crown” ranking as the world’s largest publisher in 2012, with a total revenue of $9.16 billion. Thomson’s (2005) study of university presses also emphasizes differences in economic capital. Some presses receive financial assistance from their host institutions, “ranging from annual operating grants to cover deficits to rent-free accommodation, free employee benefits, and interest-free overdraft facilities” (p. 88). Those that do not receive direct financial assistance from their host institutions expect to break even, and some that are receiving small amounts of support are experiencing growing pressure to reduce their dependence. Oxford University Press and Cambridge University Press are, as Thomson (2005) states, “in a league of their own.” At the time when Thomson’s monograph was published “OUP’s turnover year ended 31 March 2001 was £366 million ($585 million), generating a net surplus of £44 million ($70 million)” (p. 87). Compare this to a turnover for the year ended March 31, 2013, which amounts to approximately £723 million, with a net surplus of £105 million (Oxford University Press, 2013).

Symbolic capital is a criterion that is not easy to quantify in the same way as economic capital. Thomson (2005) describes this in terms of “a resource that flows back and forth between the academic field and the field of academic publishing” (p. 90). Among university presses, this particular form of capital is often aligned with institutional or host university prestige (Gump, 2006). It does however fluctuate, given that “a publisher can augment its symbolic capital by publishing authors who have gained a high reputation in the academic field, and at the same time an author can increase his or her standing in the academic field by publishing with a press that has accumulated large quantities of symbolic capital” (p. 90). The surveys carried out by Goodson et al. (1999) as well as Garand and Giles (2011) come close to addressing the issue of symbolic capital because the responses obtained from their surveyed scholars were based on publisher perceptions. We attempt to corroborate these perceptions through the use of citations, another form of symbolic capital, which is indeed measurable.

Geographical reach is the last criterion, and it is based on the number of offices that a publisher establishes worldwide. For instance, the publishing house of Palgrave MacMillan not only publishes books and journals in a wide range of subjects, but its current geographic reach is 19 countries worldwide, including North America, Europe, Asia, and Australia and New Zealand (see http://www.palgrave.com/). The geographical reach of a university press or commercial publisher is typically linked to economic capital, but it may also be historical in nature. Again, Thomson (2005) refers to Oxford University Press and Cambridge University Press because both “expanded their activities outside the U.K. in the late nineteenth and earlier twentieth centuries” when they could “take advantage of trading networks created by the British Empire.”
Digital Transformations

In addition to the publishing field’s structure, innovations are occurring in printing and publishing. As far back as the 1980s industry professionals have had a “widespread feeling . . . that digitization is bound to have a profound impact” (Thomson, 2005, p. 309). Journal publishers have been relatively quick to embrace the digital revolution, and now, with the open access movement, increasing numbers of articles are freely available for download via the Internet. Research concerning article impacts has therefore given due attention to this transformation (e.g., Moed, 2005a, 2005b, 2007). What this means, or could mean, is that book publishers are at risk of being “left behind like a beached whale” if they do not actively experiment with new technologies (Thomson, 2005, p. 332).

The prospect of producing, marketing, and selling academic books in digital form has always been attractive to publishers: “scholarly books in online environments [can] become a new source of revenue” (Thomson, 2005, p 331). The second benefit to the publisher is that the “electronic dissemination of scholarly works extricate[s] [them] from the seemingly inexorable economic logic of the field.” and, third, the electronic medium “has potential to liberate scholarly work from the constraints imposed by [print]” (pp. 331–332). Esposito (2012) aptly suggests that publishers focus on the short-form digital book, which lies somewhere in the middle between an article and a monograph.

The short form (which is really a middle form, since articles are shorter yet) has been mostly dormant for decades because the circumstances of profitable publishing in the print era could not make an economic case for the short form. Book-length work could be sold in bookstores, where it commanded a price (in today’s currency) between $12 and $25. Articles could find a place in journals or in collections of essays (sold as a book). But what to do about the 40,000-word essay? How to price it? So much of the cost of getting a book into a bookstore is in handling and transportation, so a 40,000-word piece would have to bear a price close to that of a full book. Which doesn’t make much sense if you are a customer. Thus, the short form languished—until now, when digital technology opens up new possibilities.

Several universities across the United States and elsewhere are pursuing digital strategies for promoting and selling their latest books (see the survey of the Association for American University Presses, 2014). Princeton shorts is a program at Princeton University Press, which starts with a full-length book and takes chapters or sections out to make a separate “shorter” work, in digital form (http://press.princeton.edu/PrincetonShorts/). This is one option for developing the short-form electronic book, including creating one initially from scratch. In Great Britain, Cambridge University Press has also embraced the digital movement with University Publishing Online (see http://universitypublishingonline.org/), and Mondadori, a leading publishing house in Italy, which covers 26% of the fiction and nonfiction book trade, released approximately 3,000 e-books in 2011 (Publishers Weekly, 2012).

While the publishing industry is taking a marked turn, little is known about how this is affecting the research culture of the humanities scholar. Adriaan van der Weel (2011), author of Changing our Textual Minds, indicates that the “Order of the Book is gradually disintegrating” (p. 5). The print paradigm was built upon a familiar one-way linear hierarchical order, but this new order is what he refers to now as a “digital docuverse.” Digitalization “enables new ways of accessing the text, both as a whole and as fragments,” and in terms of distribution and consumption, this “creates an entirely new relationship between author and reader” (p. 5).

In sum, we have many factors to consider when evaluating publishers. Based on the literature, it is clear how difficult it can be to address all in one study. Our objective is not to produce a definitive approach to ranking publishers but to formulate a distinct viewpoint using a unique “citation-libcitation” database. The analysis is field specific (i.e., history) and thus conveys only a small part of what makes academic publishing an intriguing enterprise. At the time this research was conducted, we did not have access to Thomson Reuters’ BKCI; hence our data-intensive approach began with an extraction of book titles cited in journal articles indexed in Elsevier Scopus.

Research Methods and Results

Data Preparation

The Scopus Journal Citation Index includes books in tagged reference lists; however, each book lacks a distinct source identification code. The researcher is forced to grapple with what is known as the reference “string” (e.g., Runge, 2005; Companion 18 Century, p. 292). Sometimes the referenced title appears in short form, and sometimes it is recorded in full, but always the publisher name is omitted. Moreover, both the author and the book title itself can be recorded inconsistently from article to article depending on the scholar who made the original citation.

In June, 2012, our project team constructed a Microsoft SQL server database, with citation records from a set of

---

1This is more often the case in Scopus than in Thomson Reuters’ Web of Science.
Elsevier Scopus journals (History and Literary Theory/Criticism) for the periods 1996–2000 and 2007–2011. The Scopus data sets were granted to us via the 2012 Elsevier Bibliometrics Research Program and comprised a total of 1,023 journals from both fields. The unique aspect of this database is that the citation records for book titles were matched and linked to publisher metadata extracted from the WorldCat® Library Catalog.2

We performed thousands of queries in WorldCat® using an API developer key. Many of the Scopus references/citations (i.e., from research articles only) could be to documents other than a book or an article (e.g., unpublished archive materials) but those preliminarily identified as a book were labeled as such because they did not have a unique Scopus source ID, a source title with a different article title, and did not contain a volume number. The data extracted from WorldCat® included the OCLC number, the ISBN of the book, the publisher name, and publishing location. With each book title classified by a new OCLC number, we also retrieved corresponding library holding counts (i.e., “libcitations”) for the Association of Research Libraries (ARL) and international or non-ARL libraries. Table 1 highlights the results of our matching procedure.

To develop a ranking of book publishers, we chose to focus on nonsourced Scopus references that were matched in WorldCat®. Sizeable counts to cited books were obtained for the 2007–2011 period for history (n = 806,985); thus our work began with this particular data set. The field itself is broadly defined on the basis of 604 different journals. This included the history and philosophy of science, the history of technology and culture, economic history, renaissance studies, medieval studies, the history of religion and biblical studies, British history, American history, Irish history, German history, Canadian history, Roman studies, African studies, and so on. The coverage of subtopics was wide ranging.

From the initial 806,985 matches, we selected only book titles that had both an OCLC number and an ISBN number. It was understood also that, to produce reliable statistics, the retrieved publisher records for all books had to be cleaned and standardized. Because we were working with combined records from two separate sources, we were highly dependent on consistencies in record keeping. Titles recorded in Scopus may have been matched to an incorrect title in WorldCat®, if one or both had been recorded incorrectly or if they referred to different books that happened to have the same title.

Data cleaning started at the level of the cited book, ensuring that all reference strings and corresponding citation counts were to individual books, followed by a process of standardizing or uniting all publisher names. The procedure included both an automated process and a manual one. For every publisher name, it was necessary to clarify, for example, that “Oxford Univ. Press” and “OUP” as well as “Oxford U. Press” were equivalent to the standardized form of “Oxford University Press.” Some press names required knowledge of the differences between particular American universities (e.g., Penn State University Press, University of Pennsylvania Press). Names that were difficult to standardize were those written in a non-English language or a non-Latin script (e.g., ROSSPEN-Rossiskaia Politicheskaia Entsiklopedia).

Many publishers were not recorded in their singular form, for instance, “Oxford University Press for the British Academy” or “Polity Press in association with Blackwell.” To simplify the outcome of this research, we decided to omit the records where a press/publisher had acted “in association with” or “in conjunction with” another or had published a book “for” another type of organization. We also chose not to standardize a publisher name if it had been altered recently by a merger (e.g., John Wiley & Sons merged with Blackwell Publishing to become Wiley-Blackwell). Instead, we kept the name as it had originally been recorded in one or more of the catalogues affiliated with WorldCat® (e.g., Blackwell Publishing). Finally, we maintained separate records of all imprints because many of these appeared in addition to their “parent” publishing name (e.g., Scribner is an imprint of Simon & Schuster).

**General Statistics**

After the data cleaning and standardization, we produced a list of 500 most highly cited publisher names from a larger set of approximately 12,000 (note: many publishing entities were cited just once). The top 500 consisted of publishers that had accumulated up to 19 citations or more in Scopus history journals during the 2007–2011 period. These 500 names were then categorized and grouped according to three

---

2WorldCat® is a union catalog that itemizes the collections of 72,000 libraries in 170 countries and territories. All libraries participate in the Online Computer Library Center (OCLC) global cooperative (see http://rlin21.rlg.org/worldcat/statistics/default.htm).
distinct types: (a) university presses; (b) commercial publishing houses; and (c) museums, libraries, special research institutes, foundations, nonprofit, and international organizations (labeled “other”).

We further categorized the publishers on the basis of their founding country. For most university presses, the label was obvious (e.g., University of Toronto Press is from Toronto, Canada), but the Internet (usually Wikipedia) was useful for verifying the geographical origin of some commercial publishers. For instance, Nauka is the name of a Russian publisher of academic books, established in 1923 in the city of Leningrad, under the original name of USSR Academy of Sciences Publisher (until 1963).

Figure 1 illustrates the number of citations and libcitations received by 500 publishers from 27 countries, including those that fit within an international category (e.g., World Health Organization; OECD—Organization for Economic Cooperation and Development). Note that the WorldCat® libcitations, or library holding counts, are much higher in number than the journal citations, yet the two variables, both skewed in distribution, are strongly correlated, with a Spearman’s rho of 0.686. Although this measure does not confirm a “cause and effect” relationship, libcitations and citations are connected; librarians are responsible for developing book collections that serve the needs of scholars. In an academic library, librarians do not necessarily “cherry-pick” books one at a time; however, the collective purchasing process still has to be strategic. Scholars want librarians to think strategically about which books will be added to the library in support of faculty research interests, just as the citation, on a micro-level, can later be selected by an author in order to support a particular argument. Strategic behavior is recognizable in both processes. Publisher catalogues, book reviews, and discussions with historians help librarians to observe research trends in subject areas, recognize curriculum needs, investigate faculty research interests, and monitor the strengths and weaknesses of collections already in place. Figure 1 indicates that publishers from the United States and Great Britain are highly favored in book-selection procedures. Clearly, the English language is dominant with respect to libcitations and citations, but certain publishers from other countries are well-catalogued and well-cited, notably those from Germany, France, and Italy.

Figure 2 shows the distributions of citations and libcitations received by three publisher categories. For university presses there is an even distribution of libcitations, but, in terms of citations, more tend to fit within the upper quartile of the distribution (i.e., above the median mark), with Cambridge University Press and Oxford University Press marked as clear outliers. The libcitation and citation distributions for the commercial publishing houses follow a similar pattern. Within the commercial set, there are also several publishing houses in competitive positions (i.e., upper 25% quartile): Routledge and Palgrave Macmillan are also clearly marked as outliers. Citation rates to alternative publishing units such as museums, institutes, and nonprofit and international organizations are not nearly as high, relative to the other publisher categories, yet still they produce a major portion of cited books (i.e., enough to include in the top 500 highly cited).

Top 50 Ranking

Reasonable numbers of international presses and commercial houses were featured in our list of 500, but, among

---

3Correlation is significant at the 0.01 level (one-tailed).
the top ranked 50 publishers, almost all originate from the United States or Great Britain. The exceptions include two German publishers and one from Switzerland. As noted in the Introduction, ranking procedures can be complicated and may lead to misleading statistics. It would be useful to employ a normalization approach to ranking based on information beyond citation counts and library holdings. Some variables, however, are difficult to control and not easily gathered from one source; for example: What are the annual sales figures for each publisher? How many books have been sold and distributed specifically for the subject of history? How many scholarly reviews have been written for the publishers’ book titles, particularly those cited during 2007–2011?

Table 2 demonstrates how citations can at least be viewed from two different perspectives. A comparison is made for 50 of the top-cited publishers on the basis of total citation counts and mean citation per book in journal articles (2007–2011). As expected, both Oxford University Press and Cambridge University Press are ranked at the top in terms of citation counts. Routledge appears to be the most prestigious commercial publisher, ranking third in the list based on total citation counts, and almost half from the list (n = 22) are commercial publishers. When we re-examine this ranking by the mean citation per book, Cambridge and Oxford move to 16th and 17th place, and the new highest ranked publishers then become Belknap, Princeton University Press, Harvard University Press, and University of North Carolina Press.

The mean citation per book is a straightforward yet rudimentary way of normalizing publisher citation counts. Several objections that have been raised against the journal impact factor also apply here. These include the inadequacy of the mean for characterizing skewed data and the fact that one cannot meaningfully compare CPBs or impact factors across fields (Leydesdorff & Bornmann, 2011). Our rankings should be interpreted with these drawbacks in mind, and in the future we may explore other procedures for normalization, such as a publisher’s proportion of most frequently cited books in a specific field. This is analogous to the percentile rank approach used in the Leiden ranking (Waltman et al., 2012).

Cambridge University Press, Oxford University Press and Routledge produce thousands of books per year for a variety of subjects and fields. Because the Scopus journal list covers a wide variety of subjects (related to history), there are ample opportunities for each of the top-cited book publishers to receive citations. This does not guarantee; however, that every printed and distributed book will have a high individual citation impact. The Belknap imprint of Harvard University Press may produce fewer books per year by comparison, but it has developed a specialist reputation in American history since 1954 when the first Harvard Guide to American History was printed. Publishing under the Belknap name thus appears to be a “prestigious” choice for some historians. Many other American university presses are at the top of our mean CPB list, but we have yet to observe the extent to which the

TABLE 2. Top 50 publishers based on total citations and citations per book (CPB) in Scopus history journals (2007–2011; commercial publishing houses printed in italics).

<table>
<thead>
<tr>
<th>Publisher name</th>
<th>Total citations counts</th>
<th>Mean CPB</th>
<th>Publisher name</th>
<th>Total citation counts</th>
<th>Mean CPB</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oxford University Press</td>
<td>6,899</td>
<td>2.972</td>
<td>Princeton University Press</td>
<td>3,176</td>
<td>4.586</td>
</tr>
<tr>
<td>Routledge</td>
<td>3,960</td>
<td>2.152</td>
<td>Harvard University Press</td>
<td>2,151</td>
<td>3.343</td>
</tr>
<tr>
<td>Princeton University Press</td>
<td>3,176</td>
<td>4.586</td>
<td>University of North Carolina Press</td>
<td>1,482</td>
<td>3.452</td>
</tr>
<tr>
<td>Palgrave Macmillan</td>
<td>2,490</td>
<td>2.097</td>
<td>University of Chicago Press</td>
<td>2,470</td>
<td>4.039</td>
</tr>
<tr>
<td>University of California Press</td>
<td>2,483</td>
<td>3.974</td>
<td>University of California Press</td>
<td>2,483</td>
<td>3.974</td>
</tr>
<tr>
<td>University of Chicago Press</td>
<td>2,470</td>
<td>4.039</td>
<td>Duke University Press</td>
<td>1,620</td>
<td>3.803</td>
</tr>
<tr>
<td>Yale University Press</td>
<td>1,862</td>
<td>3.166</td>
<td>Johns Hopkins University Press</td>
<td>1,114</td>
<td>3.397</td>
</tr>
<tr>
<td>Duke University Press</td>
<td>1,620</td>
<td>3.803</td>
<td>Cornell University Press</td>
<td>1,361</td>
<td>3.359</td>
</tr>
<tr>
<td>University of North Carolina Press</td>
<td>1,482</td>
<td>4.329</td>
<td>Basic Books</td>
<td>413</td>
<td>3.357</td>
</tr>
<tr>
<td>Cornell University Press</td>
<td>1,361</td>
<td>3.359</td>
<td>Yale University Press</td>
<td>1,862</td>
<td>3.316</td>
</tr>
<tr>
<td>Brill Academic Publishers</td>
<td>1,263</td>
<td>1.908</td>
<td>PENN, University of Pennsylvania Press</td>
<td>850</td>
<td>3.293</td>
</tr>
<tr>
<td>Stanford University Press</td>
<td>1,142</td>
<td>3.190</td>
<td>Stanford University Press</td>
<td>1,142</td>
<td>3.190</td>
</tr>
<tr>
<td>Johns Hopkins University Press</td>
<td>1,114</td>
<td>3.376</td>
<td>Penguin University Press</td>
<td>616</td>
<td>3.166</td>
</tr>
<tr>
<td>Ashgate Publishing</td>
<td>1,098</td>
<td>1.789</td>
<td>Cambridge University Press</td>
<td>7,459</td>
<td>3.118</td>
</tr>
<tr>
<td>Blackwell Publishing (now Wiley-Blackwell)</td>
<td>1,014</td>
<td>2.428</td>
<td>Oxford University Press</td>
<td>6,899</td>
<td>2.972</td>
</tr>
<tr>
<td>MIT Press</td>
<td>893</td>
<td>2.585</td>
<td>Alfred A. Knopf</td>
<td>633</td>
<td>2.913</td>
</tr>
<tr>
<td>PENN, University of Pennsylvania Press</td>
<td>850</td>
<td>3.293</td>
<td>W.W. Norton &amp; Company</td>
<td>697</td>
<td>2.797</td>
</tr>
<tr>
<td>Columbia University Press</td>
<td>844</td>
<td>2.518</td>
<td>Polity Press</td>
<td>429</td>
<td>2.719</td>
</tr>
<tr>
<td>W.W. Norton &amp; Company</td>
<td>697</td>
<td>2.797</td>
<td>Manchester University Press</td>
<td>682</td>
<td>2.520</td>
</tr>
<tr>
<td>Manchester University Press</td>
<td>682</td>
<td>2.520</td>
<td>Columbia University Press</td>
<td>844</td>
<td>2.518</td>
</tr>
<tr>
<td>Alfred A. Knopf</td>
<td>633</td>
<td>2.913</td>
<td>Berg Publishers</td>
<td>468</td>
<td>2.438</td>
</tr>
<tr>
<td>Indiana University Press</td>
<td>625</td>
<td>2.418</td>
<td>Blackwell Publishing (now Wiley-Blackwell)</td>
<td>1,014</td>
<td>2.428</td>
</tr>
<tr>
<td>Sage Publications</td>
<td>624</td>
<td>1.982</td>
<td>Indiana University Press</td>
<td>625</td>
<td>2.418</td>
</tr>
<tr>
<td>Penguin Press</td>
<td>616</td>
<td>3.166</td>
<td>University of Michigan Press</td>
<td>499</td>
<td>2.412</td>
</tr>
<tr>
<td>University of Minnesota Press</td>
<td>588</td>
<td>2.390</td>
<td>University of Minnesota Press</td>
<td>588</td>
<td>2.390</td>
</tr>
<tr>
<td>University of Michigan Press</td>
<td>499</td>
<td>2.412</td>
<td>University of Illinois Press</td>
<td>368</td>
<td>2.218</td>
</tr>
<tr>
<td>University of Toronto Press</td>
<td>487</td>
<td>2.160</td>
<td>University of Toronto Press</td>
<td>487</td>
<td>2.160</td>
</tr>
<tr>
<td>Berg Publishers</td>
<td>468</td>
<td>2.438</td>
<td>Routledge</td>
<td>3,960</td>
<td>2.152</td>
</tr>
<tr>
<td>Polity Press</td>
<td>429</td>
<td>2.719</td>
<td>Palgrave Macmillian</td>
<td>2,490</td>
<td>2.097</td>
</tr>
<tr>
<td>Basic Books</td>
<td>413</td>
<td>3.357</td>
<td>New York University Press</td>
<td>404</td>
<td>2.095</td>
</tr>
<tr>
<td>Verlag C.H. Beck</td>
<td>412</td>
<td>1.981</td>
<td>State University of New York (SUNY) Press</td>
<td>373</td>
<td>2.077</td>
</tr>
<tr>
<td>New York University Press</td>
<td>404</td>
<td>2.095</td>
<td>Vandenhoek &amp; Ruprecht</td>
<td>373</td>
<td>2.077</td>
</tr>
<tr>
<td>Viking Press (of the Penguin Group)</td>
<td>403</td>
<td>2.339</td>
<td>Boydell &amp; Brewer</td>
<td>326</td>
<td>2.073</td>
</tr>
<tr>
<td>I.B. Tauris</td>
<td>394</td>
<td>1.985</td>
<td>University of Texas Press</td>
<td>333</td>
<td>2.059</td>
</tr>
<tr>
<td>Peter Lang</td>
<td>374</td>
<td>1.280</td>
<td>Rutgers University Press</td>
<td>308</td>
<td>2.006</td>
</tr>
<tr>
<td>Vandenhoeck &amp; Ruprecht</td>
<td>373</td>
<td>2.077</td>
<td>University of Nebraska Press</td>
<td>325</td>
<td>1.994</td>
</tr>
<tr>
<td>University of Illinois Press</td>
<td>368</td>
<td>2.218</td>
<td>I.B. Tauris</td>
<td>394</td>
<td>1.985</td>
</tr>
<tr>
<td>University of Texas Press</td>
<td>333</td>
<td>2.059</td>
<td>Sage Publications</td>
<td>624</td>
<td>1.982</td>
</tr>
<tr>
<td>Boydell &amp; Brewer</td>
<td>326</td>
<td>2.073</td>
<td>Verlag C.H. Beck</td>
<td>412</td>
<td>1.981</td>
</tr>
<tr>
<td>University of Nebraska Press</td>
<td>325</td>
<td>1.994</td>
<td>Brill Academic Publishers</td>
<td>1,263</td>
<td>1.908</td>
</tr>
<tr>
<td>McGill-Queen’s University Press</td>
<td>320</td>
<td>2.048</td>
<td>Continuum Books (of Bloomsbury)</td>
<td>312</td>
<td>1.825</td>
</tr>
<tr>
<td>Continuum Books (of Bloomsbury)</td>
<td>312</td>
<td>1.825</td>
<td>Ashgate Publishing</td>
<td>1,098</td>
<td>1.789</td>
</tr>
<tr>
<td>Rutgers University Press</td>
<td>308</td>
<td>2.006</td>
<td>Peter Lang</td>
<td>374</td>
<td>1.280</td>
</tr>
</tbody>
</table>

journals indexed by Scopus focus extensively on American history compared with other topics of study.

Figure 3 compares the mean library holding counts (WorldCat®) and mean citation counts per book (Scopus) for the top 50 publishers. Here again we see a positive relationship; books by a certain publisher that are catalogued frequently in international libraries tend to receive higher citation rates in the journal literature (Spearman’s rho = 0.6144). In other words, the perceived benefit of these books for international historians corresponds significantly with their scholarly use.

*Correlation is significant at the 0.01 level (one-tailed).
Directed Journal-to-Publisher Network

With our network approach to observing publishers, we employed two mapping tools: (a) VOSViewer (Van Eck & Waltman, 2010) and (b) Pajek (de Nooy, Mrvar, & Batagelj, 2005). Both tools enabled us to explore the relationship between the Scopus journal set and the various presses/publishing houses cited by research articles. The network arcs are directed and include a selection of the top 501 strongest citation links, where the bottom threshold for link strength was set at $n=10$. In total, 354 international journals were included and 147 of the most frequently cited international book publishers. Cambridge University Press and Oxford University Press are recipients of the most in-links from this journal set ($n=354$), but significant numbers of smaller, non-British/non-American houses are featured. We spent time experimenting with the VOSViewer clustering algorithm in order to obtain interpretable results. The best option was a cluster resolution of 2.00 with a minimum size of 15. Although VOSViewer allows the user to zoom in to a specific cluster, it could not be extracted from the full map. Pajek was therefore useful for extracting and examining each cluster separately.

The VOSViewer map, shown in Figure 4, may be interpreted on the basis of subthemes and/or geographical interests. At the bottom, we see an emphasis on journals and book publishers dedicated to period studies (e.g., the Renaissance period, 16th century, 18th century, and medieval times). There are two distinct clusters to the right that emphasize the history of religion and biblical studies (e.g., early Christian studies, the study of Judaism, New Testament studies) as well as Italian and Roman studies. The top portion of the map highlights economic history, and the history and philosophy of science, and toward the left we have journals and publishers that focus on politics, diplomatic history, social history, and the Civil War in the United States. Cambridge University Press belongs to the cluster/subfield of economic history, whereas Oxford University Press is aligned with general British and Irish history, in earlier times as well as the present.

Individual analyses of the clusters point to underlying journal-to-publisher relationships. Certain presses/publishers that produce journals as well as books are directly linked. Note from Figure 5 that scholars who have written research papers for the *New Mexico Historical Review* have cited books published by the journal’s parent press, the University of New Mexico Press. Books published by the University of California Press have often been cited in research papers from the journal *Agricultural History* (i.e., also from the same “parent” publisher). Figure 6 illustrates the importance of Brill Academic Publishing for topics relevant to the history of religion. Books published by both Brill and Walter de Gruyter have also been cited frequently in research articles from their own journals.

Although it is possible to rank presses/publishers on one indicator alone, our network approach to mapping directed citations suggests that it may be prudent to think in terms of specialization. Historians do not necessarily have to publish with a high-ranking press such as Cambridge or Oxford to gain recognition by citation. The choice of publisher also depends on the type of academic audience one would like to reach. For example, a scholar of Latin American history/Hispanic studies might publish with the University of California Press or Duke University Press, and the historian who wants to make an impact with new research concerning the history of Christianity might choose to publish a book with Brill.

Discussion

Can we rank scholarly publishers? Without access to a complete index of cited books, it is possible, but the
FIG. 4. Three hundred fifty-four Scopus history journals with directed citation links to 147 book publishers (n = 16 clusters). [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]
procedure is difficult. However, the amalgamation of two bibliographic resources, Scopus and WorldCat®, following a detailed matching, cleaning, and standardizing procedure for publisher names and book titles, has afforded us insights that we might otherwise not have gained had we been granted the “perfect” index. This work has generated some results that were expected, along with information that tedious data handling has brought to light. findings reflect
Commercial organizations such as Thomson Reuters and Elsevier are now faced with the question: “To include or not to include?” Book metadata are currently being added to the BKCI and Scopus, but little is known about the selection criteria for titles and publishers. Certain choices could be detrimental not only for the publishing industry but for historians as well, particularly those with national or regional interests. Our journal-to-publisher network confirms that it is still a standard practice for some scholars to publish research articles and books in French, German, and Italian. A selection of non-English-language books, as well as those originally printed in non-Latin scripts may or may not be added to the new indices depending on how much effort is made to achieve comprehensiveness. Will this continue in a climate of financial cutbacks, internationalism, and digital transformations? In terms of writing and publishing, historians are divided. Some are willing to forego local/regional interests, write as much as possible in English (if not in an English-speaking country), and embrace the notion of high-ranking international journals. Others are less willing to publish one or more journal articles versus a full-length monograph, because short publishing strategies compromise the in-depth treatment of their subject. Now, however, the electronic short-form book could actually change the scholar’s outlook. The process of ranking book publishers could alter promotion and tenure expectations, but new departmental expectations for producing research might also change how we value different forms of text when establishing publisher rankings.

This study was not designed to provide historians with a definitive list of publishers that they should now choose from to increase their scholarly impact. It was an experiment, and our experimental approach has led us to the following conclusions. Citation indices that include books should establish clear records of presses and commercial publishers when they publish independently or in conjunction with other organizations. The cleaning and standardizing process for this study showed that many university presses, for example, act in association with other entities such as museums, libraries, and special research institutes. Moreover, several institutes, organizations, museums, and libraries seem to have established themselves as publishing units in their own right. When publisher names are added to the new indices, especially commercial publishers, they should be recorded consistently and imprints need to be distinctly identified as well, since imprints do not necessarily have the same role as their ‘parent’ publishing houses. A book publisher that establishes a reputation as a topic specialist is not in the same league as an international, comprehensive publisher. Clearly, the regional specialist will not benefit from being evaluated or ranked alongside larger competitors with more economic capital. Comprehensive presses/publishing houses will naturally attract more citations simply by having distributed more books. However, the book publisher that specializes in a regional topic need not aim to be comprehensive; there are enough presses and commercial houses that share similar strengths and have survived well enough for a comparative evaluation, despite the putative publishing crisis.

Acknowledgments

The authors are grateful to both the Elsevier Bibliometrics Research Programme (http://ebrp.elsevier.com/) and OCLC WorldCat® for providing the data sets that were used to create the unique database for this study. We also thank the Center for Digital Humanities in Amsterdam for financing our project, as well as Maurits van Bellen and Robert Iepsma for assisting us with the data-cleaning and standardization procedure.

References


Calver, M.C., & Bradley, J.S. (2009). Should we use the mean citations per paper to summarise a journal’s impact or to rank journals in the same field? Scientometrics, 81(3), 611–615.


Garfield, E. (1972). Citation analysis as a tool in journal evaluation. Journals can be ranked by frequency and impact of citations for science policy studies. Science, 178(4060), 471–479.


Leydesdorff, L., & Felt, U. (2013). Edited volumes, monographs and book chapters in the Book Citation Index (BKCI) and Science Citation Index (SCI, SoSCI, A&HCI). Journal of Scientometric Research, 1(1), 28–34.


Thomson Reuters. (2013). Putting books back into the library: Completing the research picture. The Book Citation IndexSM. Retrieved from http://wokinfo.com/products_tools/multidisciplinary/bookcitationindex/

