Analyzing the citation characteristics of books: edited books, book series and publisher types in the book citation index

Daniel Torres-Salinas · Nicolás Robinson-García · Álvaro Cabezas-Clavijo · Evaristo Jiménez-Contreras

Abstract This paper presents a first approach to analyzing the factors that determine the citation characteristics of books. For this we use the Thomson Reuters' book citation index, a novel multidisciplinary database launched in 2011 which offers bibliometric data on books. We analyze three possible factors which are considered to affect the citation impact of books: the presence of editors, the inclusion in series and the type of publisher. Also, we focus on highly cited books to see if these factors may affect them as well. We considered as highly cited books, those in the top 5 % of those most highly cited in the database. We define these three aspects and present results for four major scientific areas in order to identify differences by area (science, engineering and technology, social sciences and arts and humanities). Finally, we report differences for edited books and publisher type, however books included in series showed higher impact in two areas.

Keywords Book citation index · Monographs · Highly cited · Publishers · Databases

D. Torres-Salinas
EC3: Evaluación de la Ciencia y de la Comunicación Científica, Centro de Investigación Biomédica Aplicada, Universidad de Navarra, Madrid, Spain
e-mail: torressalinas@gmail.com

N. Robinson-García (✉) · Á. Cabezas-Clavijo · E. Jiménez-Contreras
EC3: Evaluación de la Ciencia y de la Comunicación Científica, Departamento de Información y Documentación, Universidad de Granada, Granada, Spain
e-mail: elrobin@ugr.es

Á. Cabezas-Clavijo
e-mail: acabezasclavijo@gmail.com

E. Jiménez-Contreras
e-mail: evaristo@ugr.es

🖄 Springer

Introduction

One of the basic outcomes from the field of bibliometrics and citation analysis is the characterization of document types and field-based impact which allows fair comparisons and a better understanding of the citation patterns of researchers (Bar-Ilan 2008). These studies are of great relevance in the field as they contextualize the impact of research as well as 'anomalies' such as the higher impact of reviews or editorial material versus research papers (Archambault and Larivière 2009), the impact of research collaboration (Lambiotte and Panzarasa 2009), or the importance of monographs in the Humanities (Hicks 2004). In this context, the role played by citation indexes in general and by those developed by Eugene Garfield and constructed by Thomson Reuters in particular, have been of key importance in developing these analyses (Garfield 2009). However, these citations indexes are mainly devoted to scientific journals and neglect other communication channels such as monographs. In this respect, Campanario et al. (2011) analyze the effect different document types have on journal impact factors.

Many attempts have been made to analyze the impact of books using not only citation analysis but alternative metrics as well as databases. White et al. (2009) developed the concept of 'libcitation' in which they conceive library holdings as a measure of the impact of books and, hence, use this indicator to compare research units. Following this line of thought, Linmans (2010) suggests the use of library holding analysis to complement citation analysis when evaluating research units, especially in the social sciences and the humanities. Torres-Salinas and Moed (2009) studied this idea in depth, exploring what they termed library catalog analysis, which seeks to develop a bibliometric methodology based on the presence of books in online catalogues rather than in citations. Finally, Zuccala and van Leeuwen (2011) explored the potential of including book reviews in bibliometric research evaluations for the humanities.

Regarding the use of citation analysis, Hammarfelt (2011) explores the possibility of using references from journal articles from the web of science to books, which are non-citable items, in order to analyze highly cited monographs. He highlights the difficulties this methodology entails when focusing on the impact of specific research units. Kousha et al. (2011) go a step further and explore the use of Google books and Google scholar as alternative databases for retrieving, not only citations from journal articles to books, but also from books to books, showing a greater coverage than other databases such as Scopus.

Despite all these efforts, little is known of the characterization of book citation patterns. What is more, most of these studies are based on small data sets and specific disciplines. For instance, Cronin et al. (1997) compared a list of highly influential authors derived from journal citations with another derived from books in the field of Sociology, concluding that these two publication types reflect complementary aspects of a fragmented picture. Tang (2008) took a step further and investigated the citation characteristics of a sample of 750 monographs in the fields of religion, history, psychology, economics, mathematics and physics, finding significant differences when compared with findings in the literature on citation in journal articles. Georgas and Cullars (2005) adopted a different approach and analyzed the citation characteristics of literature in Linguistics to see if researchers' habits in this field were more closely related to the social sciences than the humanities. In general, the conclusions of these studies must always be taken with caution as they cannot be extended to all scientific fields.

The launch in 2011 of the Thomson Reuters' book citation index (henceforth BKCI)—which provides large sets of bibliometric data on monographs—was intended to resolve issues about data availability and may serve as a resource for further citation analyses and a

more thorough understanding of the citation characteristics of books. It covers scientific literature since 1999 and, as with the other citation indexes, follows a rigorous selection process using the following principle criteria (Testa 2010): currency of publications, complete bibliographic information for all cited references, and the implementation of a peer review process. Despite being very recent, the BKCI has already been the object of study. Indeed, many studies analyze its coverage, caveats and limitations. Leydesdorff and Felt (2012) analyze citation differences between monographs, edited volumes and book chapters by comparison with journal articles. They highlight the conceptual limitations that arise when book chapters are considered to be individual contributions and the manner in which this may affect any subsequent analysis of researcher output; the scarce number of citations that books receive; and the problems that arise from ignoring differences between book series and annual series.

The inflation of publication counts is reported by Gorraiz et al. (2013) who also point to limitations such as the lack of cumulative citation counts or the absence of affiliation data. Other studies of the BKCI are those reported by Torres-Salinas et al. (2012, 2013a, b). Here a different perspective is taken, analyzing the database from the publishers' viewpoint: the field-specialization of publishers, concentration, citation patterns of book chapters and even suggesting the possibility of reproducing the Journal Citation Reports model by developing Book Publisher Citation Reports (Torres-Salinas et al. 2012). The present paper seeks to explore further aspects of books as a channel of scholarly communication and the citation characteristics that define them, by making use of the BKCI. This study aims to compare the citation patterns of different book types. To do so, we define these in terms of three variables and present results for each of four major macro-areas of scientific knowledge (Science, Social Science, Engineering & Technology and Humanities). Specifically, our objective is to answer the following research questions:

(1) Edited books versus non-edited books. There is a perception that edited books usually have a greater impact than non-edited books (Leydesdorff and Felt 2012). In this study we aim to determine the extent to which this is true. Do differences exist between areas of knowledge?

(2) Series books versus non-series books. The prestige or impact derived from the series in which the book is included could be considered evidence of the quality books may have due to the editorial process they have undergone in order to be published. Is there any empirical evidence for this claim in terms of higher citation rates?

(3) Publisher type. Is the publishers' prestige related to a book's impact? Which publishers receive more citations: university presses, commercial publishers or other academic publishers?

Also, we take special interest in highly cited books. That is, the 5 % of most cited books in each specific area. We try to determine the characteristics of this special subset of books in terms of our previously-defined variables.

Materials and methods

This section is structured as follows. First, we describe the data retrieval and processing procedures, indicating the normalization problems encountered and how these were solved. In "Overview and main limitations of the book citation index" subsection, we define the areas under study and how these were constructed, basing our methodology on previous studies and offering an overview of the distribution of books by area in the BKCI as well as

highlighting the main aspects of this database which must be taken into account when analyzing the results of our study. Then, in "Definition of variables and indicators" sub-section, we define the variables analyzed and describe the methodology followed as well as the statistical analysis undertaken in order to pursue the goals of the study.

Data retrieval and processing, and definition of areas

Although this study is centered solely on books, due to the way the database is designed, we downloaded all records indexed as 'book chapter' and as 'book' in the BKCI. The reason for doing this is derived from the way the BKCI is designed and it means some modifications to the citation count are needed in order to overcome previously-mentioned limitations that we will now discuss. Book chapters are considered separate records; which means that publication counts are inflated, which benefits books with more chapters. Consequently, only records designated as books were included in the study. However, we cannot rely solely on the information provided by the records indexed as 'book' when considering citations. As mentioned above, citation counts are not cumulative so the citation data offered by the database for a given book does not contain those citations received for the book chapters. Hence, in order to obtain the total number of citations received by each book we must calculate the total number of citations received by its chapters.

The download took place in May 2012 and the study period was 2005–2011. The chosen time period is based on the availability of data at the time of retrieval. Data was entered into a purpose-built relational database. During data processing, publisher names were normalized as many had variants that differed depending on the location of their head offices in each country. For instance, Springer uses variants such as Springer-Verlag Wien, Springer-Verlag Tokyo, Springer publishing co, among others. Note that a fixed citation window was used, which means older books are more likely to be cited than others. Also, we emphasize the fact that citations included in the BKCI come from all the citation indexes provided by Thomson Reuters (SCI, SSCI and A&HCI) and not only the BKCI. Once the total number of book citations had been established we excluded Annual Reviews. This publisher, which includes a total of 234 records, was found to include journals rather than books, as noted by Torres-Salinas et al. (2013a) who observe that the citation pattern followed by this publisher is not comparable with that of the rest of the database. The final data set consisted of 28,634 books.

Overview and main limitations of the book citation index

In order to provide the reader with a general overview, we decided to cluster all BKCI subject categories (249) into four macro-areas: Arts & Humanities (HUM), Science (SCI), Social Sciences (SOC) and Engineering & Technology (ENG). Aggregating subject categories is a classical perspective followed in many bibliometric studies when adopting a macro-level approach (Moed 2005; Leydesdorff and Rafols 2009). These aggregations are needed in order to provide the reader with an overview of the whole database. Thus, we minimized the chances of overlapping for records assigned to more than one subject category. Also, we consider that these areas are easily identifiable by the reader as they establish an analogy with the other Thomson Reuters' citation indexes (Science Citation Index, Social Science Citation Index and Arts & Humanities Citation Index). With the exception of science which, due to the heterogeneity of such a broad area, was divided into two: Science and Engineering & Technology. In Table 1 we show the distribution of the sample of books analyzed across the four areas.

Table 1 Distribution of books analyzed in this study by areas as well as total and average citations received according to the BKCI 2005–2011

Discipline	Acronym	Total books	Total chapters	Average chapters by book	Books (%)	Citations	Average citations
Engineering & Technology	ENG	3,871	49,076	12.68	14 %	34,705	8.97
Arts & Humanities	HUM	8,251	94,825	11.49	29	52,224	6.33
Science	SCI	9,682	137,027	14.15	34 %	241,230	24.91
Social Science	SOC	10,637	129,754	12.20	37 %	99,943	9.40
Total books without duplicates		28,634	363,152	13.25	100 %	392,429	13.70

Finally, we underline the main limitations of the BKCI. As a recent product, we consider it necessary to summarise the key aspects indicated in previous studies as the coverage and characteristics of this database may well influence the results derived from this study. Hence these aspects should be borne in mind when observing the results derived from our study. These can be summarized as follows:

– *Language bias* The BKCI has a strong bias towards english speaking countries (96 % of the records in the database are in english language) and specifically towards the United States and the United Kingdom as place of publication (these two countries represent 75 % of the database). This seriously affects certain fields especially in the social sciences and the humanities which are characterized by a strong national orientation (Hicks 2004).
– *Great concentration of publishers* The database contains records from up to 18 publishers, with a high presence of commercial publishers to the detriment of university presses and other academic publishers. Also, three publishers (Springer, Palgrave and Routledge) represent half of the database.
– *Inclusion of Annual Reviews as books* The BKCI includes Annual Reviews with their ISSN considering them books. This generates a considerable distortion in terms of citations because reviews tend to accumulate a greater number of citations.
– *Inflation on publication counts* As discussed in the Introduction, the BKCI considers book chapters and books to be independent records, meaning that publication counts may be misleading as this suggests that authors of books with a large number of chapters have a greater number of publications than those with fewer chapters (for further discussion of this issue the reader is referred to Leydesdorff and Felt 2012).
– *Dispersion of citations* Due to the distinction between books and book chapters, citations to each of them are also considered as independent. This means that citations received by book chapters are not included in the citation counts of books, therefore these must be added when working solely with books.
– *Citation errors* Torres-Salinas et al. (2013b) and Gorraiz et al. (2013) reported errors in the citation count when comparing those included in the BKCI with those found using the Cited Reference Search option.
– *Conceptual problems* Unlike journal articles, books present many issues which must be resolved before treating them for bibliometric purposes. For instance, should we consider new editions of a book as separate records? Or, must translations of a book be included in the same record? Depending on the solution reached this may lead to new

technical problems such as assigning more than one publisher to a record (if translations are carried out by different publishers).

Definition of variables and indicators

Here we define and describe the three variables analyzed to characterize book citations: presence of editors, inclusion of books in a series and publisher type.

Presence of editors

In order to analyze edited and non-edited books we considered as the former those which were indexed as such according to the Book Editor (ED) field provided by the BKCI. We considered non-edited books those which had no information in this field. For instance, the book entitled 'Power Laws in the Information Production Process: Lotkaian Informetrics' which is single-authored by L. Egghe has no information in the ED field, therefore it is considered as a non-edited book. In contrast, the book 'Web 2.0 and Libraries: Impacts, Technologies and Trends' is edited by D. Parkes and G. Walton and has contributions from different authors, therefore it is considered an edited book.

Inclusion in a series

In order to analyze the inclusion of books in a series we used the field defined in the BKCI as Series (SE), tagging as such those records which contained information in this field and as non-series, those which did not. We identified a total of 3,374 different series in the BKCI. The series with a higher number of books indexed in the BKCI for each field are: 'Studies in Computational Intelligence' published by Springer (243 books) for engineering and technology, 'New Middle Ages' by Palgrave (49 books) in Arts & Humanities; 'Methods in Molecular Biology' by Humana Press Inc (232 books) in science, and 'Chandos Information Professional Series' by Chandos (118) in social sciences.

Publisher type

After data processing and normalizing publisher names as mentioned above, 280 publishers were identified in the BKCI. These publishers were then divided into the following three categories:

– *University press* Defined as any publisher belonging to a university such as the Imperial college press or Duke university press.
– *Non-university academic publisher* Publishers belonging or related to organizations such as research institutions, scientific societies or any other type of entity not linked to universities such as the Royal Society of Chemistry or the Technical Research Centre Finland.
– *Commercial publisher* Publishers considered in this group are those not related to universities or any other scientific entity but to profit-oriented firms such as Routledge or Elsevier.

Finally, we characterized the factors that determine book citations. We report the effect size of citations for each variable and by area as a means of analyzing the extent of citation

differences. For this, we use Cohen's *d* and *r*, and the criteria proposed by Cohen (1988). Hence, effect size is considered small, medium or large according to the following values:

$$r \text{ effects}: \text{ small } \geq 0.10 \text{ medium } \geq 0.30 \text{ large } \geq 0.50$$
$$d \text{ effects}: \text{ small } \geq 0.20 \text{ medium } \geq 0.50 \text{ large } \geq 0.80$$

Furthermore, we analyzed the characteristics of highly cited books (henceforth HCB), that is, the 5 % most highly cited for each of the four macro-areas under study. We identified 1,534 books as HCB.

Results

In this section we report the results of the study according to the three variables analyzed. Firstly, we offer the results for presence of editors; secondly, we show those for the inclusion of books in series; and thirdly, we focus on publisher type. In the fourth sub-section, we report the effect sizes for each variable. Finally we characterize the top 5 % of HCB in terms of these variables.

Edited versus non-edited books

In Table 2 we offer an overview of the sample of books analyzed according to the presence of editors. Overall, from the total sample (ALL), 12,646 books (44 %) had been edited while 15,988 books (56 %) had not. Edited books have a significantly higher citation rate than those which are non-edited, as shown by the average and median values. This occurs in the four areas studied. The most significant differences are found in the area of Science where edited books have an average of 35.41 citations per book as opposed to 10.16 citations per non-edited book. Also, edited books have a higher average of chapters per book and reach higher citation values as indicated by the standard deviation and median values. To a lesser extent, this also occurs in the social science and engineering and technology areas. The lowest differences between edited and non-edited books are found in the area of arts and humanities, where edited books have a citation average of 7.61, while non-edited books have an average of 5.81.

Inclusion in series versus non-inclusion in series

There are a total of 17,789 books included in series (62 % of the total share) while those not included in series amount to 10,845 (38 %) (Table 3). The distribution of books in series varies according to the area. Science and engineering and technology are the areas with the highest shares, especially the latter where books in series represent 71 %. With regard to citation average and median values of books included in series, these two areas—and especially science—show the most significant differences. In contrast, there are no significant differences in the Social Sciences, and the median value for both included and non-included books, is 3.00. The only exception noted is in Arts & Humanities, where non-included books have a higher citation average and median value than those included in series. Regarding the average number of book chapters by book, books not included in series have a higher average number of chapters, however, when observing each area we see that this is the case only for science, whereas there are almost no differences in the other areas.

Table 2 Citation and statistical indicators

Discipline	Type	Books	Books (%)	Average chapters by book	Citation average	SD	Median
ALL	Edited books	12,646	44	16.32	21.81	±99.35	5.00
	Non edited books	15,988	56	9.81	7.16	±7.61	2.00
ENG	Edited books	1,841	48	15.93	12.00	±24.59	4.00
	Non edited books	2,030	52	9.71	6.21	±15.82	1.00
HUM	Edited Books	2,384	29	16.20	7.61	±15.26	3.00
	Non edited books	5,867	71	9.59	5.81	±14.45	2.00
SCI	Edited books	5,658	58	16.97	35.41	±145.45	7.00
	Non edited books	4,024	42	10.30	10.16	±35.96	2.00
SOC	Edited books	4,254	40	15.56	12.0	±29.24	4.00
	Non edited books	6,383	60	9.96	7.66	±24.35	2.00

Edited versus non-edited books 2005–2011

Table 3 Citation and statistical indicators

Discipline	Type	Books	Books (%)	Average chapters by book	Citation average	SD	Median
ALL	Series books	17,789	62	13.41	12.62	±45.68	3.00
	Non series books	10,845	38	15.88	10.98	±95.38	3.00
ENG	Series books	2,746	71	12.59	10.06	±28.25	3.00
	Non series books	1,125	29	12.86	7.6	±23.50	2.00
HUM	Series Books	4,585	56	11.24	5.91	±14.43	2.00
	Non series books	3,666	44	11.83	6.86	±15.04	3.00
SCI	Series books	6,349	66	12.92	29.63	±69.19	5.00
	Non series books	3,333	34	16.54	15.93	±169.37	2.00
SOC	Series books	5,854	55	12.03	9.1	±27.10	3.00
	Non series books	4,783	45	12.40	9.75	±25.76	3.00

Included in series versus non-included in series books 2005–2011

Type of publisher

Overall, 83 % of the books included in the BKCI belong to commercial publishers, followed at a considerable distance by university presses (14 %) and non-university academic publishers (3 %) (Table 4). This distribution varies substantially depending on the area. In engineering and technology the presence of commercial publishers is even higher (97 %), while in the arts and humanities and the social sciences, the university presses have a higher presence (27 and 15 %, respectively).

When analyzing citation average and median values in general, we find a pattern common to all areas: the commercial publishers are those with the lowest citation averages and the university presses are those with the highest number of citations. The highest difference is noted in the arts and humanities, where the latter show a citation average of 11.62 while the former have values of 4.32. The same is true of the Social Sciences, where

Table 4 Citation and statistical indicators

Discipline	Type	Books	Books (%)	Average chapters by book	Citation average	SD	Median
ALL	Academic non univ	919	3	13.60	23.90	53.40	5.00
	Commercial publisher	23,843	83	14.51	12.36	39.67	2.00
	University press	3,872	14	13.51	20.22	156.60	7.00
ENG	Academic non univ	72	2	14.67	16.22	±28.45	5.00
	Commercial publisher	3,726	97	12.61	8.62	±18.90	2.00
	University press	57	1	13.84	23.96	±65.73	7.00
HUM	Academic non univ	51	1	13.47	5.33	±9.55	2.00
	Commercial publisher	5,906	71	11.70	4.32	±9.86	2.00
	University press	2,270	28	10.92	11.62	±22.20	6.00
SCI	Academic non Univ	696	7	13.51	28.26	±59.75	6.00
	Commercial publisher	8,375	87	14.16	22.81	±61.23	3.00
	University press	517	6	15.35	50.88	±421.40	9.00
SOC	Academic non univ	181	2	7.44	10.71	±24.24	3.00
	Commercial publisher	8,816	83	12.43	7.33	±21.13	2.00
	University press	1,626	15	11.50	20.40	±44.17	8.00

Included in series versus non-included in series books 2005–2011

university presses have a citation average of 20.40 versus commercial publishers, with 7.33. These differences are also significant for engineering and technology and Science, although not to the same extent. We do not observe a clear pattern for each field in terms of the number of book chapters per book.

Effect sizes

In Table 5 we report the effect sizes between the groups analyzed for each of the three variables under study. We include Cohen's d as well as correlation r and use the convention proposed by Cohen to interpret the extent of the effect size, as mentioned earlier. As observed, for most groups, the effect size of the differences on their citation characteristics is small, with d never reaching a value of 0.3. We only find medium effect sizes ($d \geq 0.30$) when analyzing differences by publisher type. In engineering and technology, we observe a medium effect size between academic and commercial publishers ($d = 0.31$). In the areas of arts and humanities there is a medium effect size between academic publishers and university presses ($d = -0.37$) and between commercial publishers and university presses ($d = -0.43$). Finally, in social sciences we observe a medium effect size between commercial publishers and university presses ($d = -0.38$).

Citation characteristics of highly cited books

In Fig. 1, we include the citation characteristics of the HCB for the four macro-areas under study according to the presence of editors, inclusion in series and publisher type. Of 1,534 HCB, 65 % were edited while 35 % were non-edited. This general pattern also appears in three of these areas, especially in Science with 90 % of HCB being edited books, followed by engineering and technology (65 %) and Social Sciences (57 %). The only exception is

Table 5 Effect sizes between groups for each of the variables under study: presence of editors, inclusion in series and type of publisher

Discipline	Type	d	r	Effect size
ALL	Edited versus non edited	0.21	0.10	Small
	Series versus non series	0.02	0.01	Small
	Academic versus commercial	0.25	0.12	Small
	Academic versus university press	0.03	0.02	Small
	Commercial versus university Press	−0.01	−0.03	Small
ENG	Edited versus non edited	0.28	0.14	Small
	Series versus non series	0.09	0.05	Small
	Academic versus commercial	0.31	0.16	Medium
	Academic versus university press	−0.15	−0.08	Small
	Commercial versus university press	−0.32	−0.16	Small
HUM	Edited versus non edited	0.12	0.06	Small
	Series versus non Series	−0.06	−0.03	Small
	Academic versus commercial	0.10	0.05	Small
	Academic versus university press	−0.37	−0.18	Medium
	Commercial versus university press	−0.43	−0.21	Medium
SCI	Edited versus non edited	0.24	0.12	Small
	Series versus non series	0.11	0.05	Small
	Academic versus commercial	0.09	0.04	Small
	Academic versus university press	−0.08	−0.04	Small
	Commercial versus university press	−0.09	−0.05	Small
SOC	Edited versus non edited	0.16	0.08	Small
	Series versus non series	−0.02	−0.01	Small
	Academic versus commercial	0.15	0.07	Small
	Academic versus university press	−0.27	−0.13	Small
	Commercial versus university press	−0.38	−0.19	Medium

found in arts and humanities where the percentage of edited HCB is lower than that for non-edited with 42 % of the total share of HCB. However, when interpreting the data for HCB presented in Fig. 1, the total share of the BKCI (included in Table 2) must be taken into account. For instance, in engineering and technology there is a higher share of edited books. In the case of the arts and humanities, edited books represent only 28 % of the total share. However, the share for HCB edited is 42 % of the total HCB in this area. This means that HCB are more commonly edited than non-edited books.

If we focus on HCB included in series, we observe that 65 % of most cited books overall belonged to a series. Also, books included in series are better represented in the area of engineering and technology, while in science the proportion is 50 %. In social sciences and arts and humanities, the HCB with greater presence are those not included in series. In the four areas analyzed there are no significant differences in the overall distribution of books included and non-included in series as well as in the distribution of HCB. Therefore, while 71 % of books in engineering and technology are included in series, 67 % of the HCB in this area are also included in series, following a similar distribution.
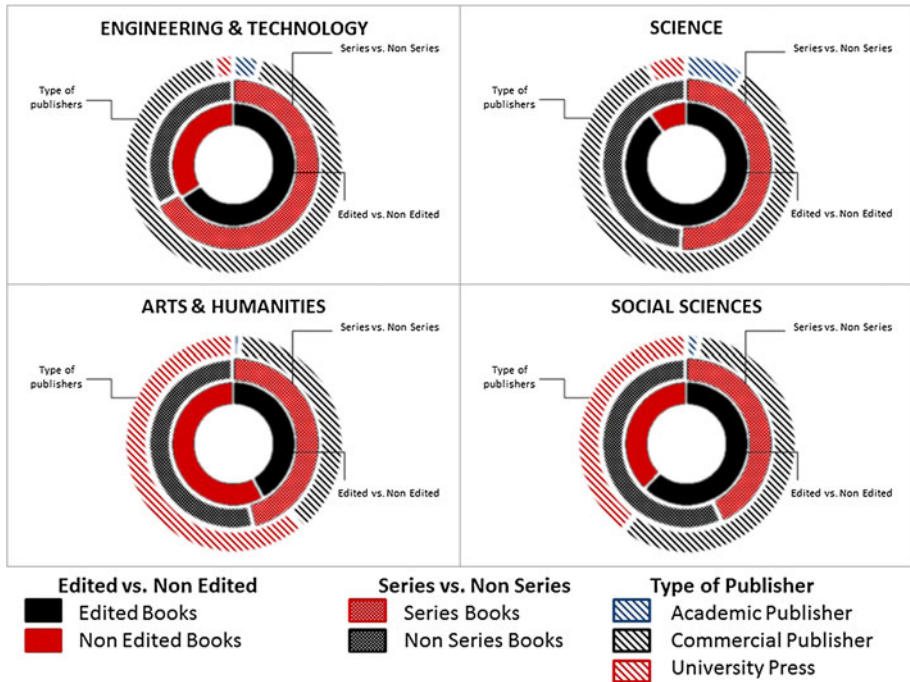
Fig. 1 Citation characteristics of the 5 % most highly cited books in the BKCI for the areas of engineering and technology, arts and humanities, science and social sciences. 2005–2011 time period

Regarding the distribution of HCB according to the publisher type, the most significant result is the high representation of HCB among books from university presses, almost always higher than the other two publisher types for all areas. For instance, for all areas university presses represent 14 % of the total share. However, when focusing on HCB, they represent 30 %. This phenomenon is especially relevant in two of the four areas under study. Thus, in Arts & Humanities 28 % of the total share are published by university presses, but 61 % of the HCB are from this type of publisher. Similarly, in the Social Sciences they represent 15 % of the total share but have 40 % of the total HCB.

Discussion and concluding remarks

This paper analyzes the citation characteristics of books according to three variables: the presence of editors, their inclusion in series and the type of publisher. For this, we used a sample of 28,634 books indexed in the BKCI for four macro-areas during the 2005–2011 time period. We emphasize that the BKCI is a novel database constructed by Thomson Reuters which opens new opportunities for analyzing citation phenomena in books as it currently occurs with journals, where these characteristics have already been thoroughly analyzed (e.g., Peritz 1981; Aksnes 2003). By comparison with journal articles, many issues remain unsolved when analyzing books and their different typologies (i.e., edited, non-edited, included in series, etc.) for bibliometric purposes. Just as we know that review articles tend to receive more citations than research articles or that editorial material

receives few citations, we still need to improve our understanding of the way the type of book affects their citation pattern. In this paper, we explore this issue through a large multidisciplinary dataset divided into broad research areas.

Our findings suggest that there are differences according to these factors. Despite acknowledging their importance, especially in the fields of the Social Science and Humanities, the bibliometric community has lacked data sources that would allow us to retrieve relevant citation and publication data. The emergence of Google Books or the BKCI has made this possible. But, before embracing them and developing bibliometric indicators for research evaluation purposes just as we have with journal articles we must fully analyze and understand their citation characteristics in order to adjust and adapt them correctly to our bibliometric toolbox (Leydesdorff, 2009). This is important especially at a time when efforts are being redoubled in order to obtain bibliometric measures regarding the research evaluation of books. Despite finding differences when comparing types of book, the effect size of these differences is not as large as would be expected (Table 5), this may be due to the coverage of the database which has been reported as having serious limitations (Torres-Salinas et al. 2013b). In order to test the extent of these differences, further research is needed; focusing on specific fields. Furthermore, more thorough analysis of the conceptual differences between edited and non-edited books is needed. Should they receive a different treatment when being subjected to bibliometric analysis? Should we descend to the level of book chapter only when dealing with edited books?

Another issue to consider has to do with the consequences and possibility of applying publication and citation analysis to books for research evaluation purposes. In contrast to findings reported by Gorraiz et al. (2013), our data does include affiliation information which could be used to analyze research units such as countries or institutions. This difference may be due to the time gap between the downloading of the respective studies. Torres-Salinas et al. (2012) suggested developing rankings of publishers according to their citation data. In this study, the differences when larger effect sizes were found related to comparisons between publisher types. Perhaps this should be taken into account before attempting this goal. Table 6 shows the main findings of the study.

(1) Non-edited books are more common in the arts and humanities and the social sciences than edited books. This could lead us to presume that non-edited books would be better more cited. However, edited books have a greater impact than non-edited books in all areas. This may be due to the effects of working collectively with a more diversified content and therefore, more chances of being cited. Another possible explanation may lie in the average number of book chapters per book. As observed in Table 2, edited books have on average a much higher number of chapters; that is citable items, as conceived by the BKCI. This may enhance their opportunity of receiving more citations than non-edited books.

(2) The inclusion of books in series is more frequent in the areas of engineering and technology and science, while in the arts and humanities and social sciences, their distribution is more homogeneous. However, the impact of books according to their inclusion in series varies depending on the area. In engineering and technology and science, books included in series receive more citations than those which are not, although this is not as significant as in other cases. In the arts and humanities, books not included in series are those with higher impact, but there are fewer differences. In the case of the social sciences, the differences are almost negligible.

(3) Considering publisher type, most books indexed in the BKCI belong to commercial publishers, especially in the areas of engineering and technology and science. Though

Table 6 Highlights of the main findings of this study analyzing the factors which determine the citation of books in four major areas

	ENG	HUM	SCI	SOC
Characteristics of the BKCI coverage				
Edited versus non edited	Edited and non-edited books are equally distributed	There are more non-edited books than edited (71 %)	Edited and non-edited books are not equally distributed	There are more non-edited books than edited (60 %)
Series versus non series	Most books are included in series (71 %)	Books included and not included in series are equally distributed	Most books are included in series (66 %)	Books included and not included in series are not equally distributed
Type of publisher	Most books are from commercial publishers (97 %)	Most books are from commercial publishers (71 %)	Most books are from commercial publishers (87 %)	Most books are from commercial publishers (83 %)
Citation characteristics of books included in the BKCI				
Edited versus non edited	Edited books are more cited	Edited books are more cited	Edited books are more cited	Edited books are more cited
Series versus non series	Books included in series are more cited	Books not included in series are more cited	Books included in series are more cited	There are no citation differences
Type of publisher	Books from university presses are more cited	Books from university presses are more cited	Books from university presses are more cited	Books from university presses are cited

Data BKCI 2005–2011

the distribution is similar in the arts and humanities and the social sciences, the university presses are better represented in these areas. However, one phenomenon is common across all areas: books published by university presses receive significantly more citations than the others. At this point we must approach this statement with caution as, after examining these publishers, we find that the university presses included in the BKCI are considered of huge prestige such as Cambridge UP, Princeton UP and University of California P. That is, books from university presses may be highly cited due to the better selection process of books and topics than that followed by commercial publishers, such as Elsevier, Routledge or Palgrave, for instance.

Finally, we must point out that the results offered in this analysis inherit the shortcomings of the database from which the data was retrieved. The BKCI is an on-going project which still shows significant limitations. Some of these may affect the results presented such as a bias towards english language publications (96 % of its books are written in this language and 75 % of the publishers come from the United Kingdom or the United States) and a great concentration of publishers. For example, Springer, Palgrave and Routledge alone account for 50 % of the total database. Therefore, these issues must be taken into consideration when analyzing our findings. However, the large data set used may be a significant step towards a better comprehension of the citation characteristics of books. Also, this study presents a global overview of the scientific knowledge, dividing it into broad areas. In order to deepen on the citation characteristics of books further analyses are needed focusing on specific disciplines, especially from the fields of social sciences and humanities.

References

Aksnes, D. W. (2003). Characteristics of highly cited papers. *Research Evaluation, 12*(3), 159–170.

Archambault, E., & Larivière, V. (2009). History of the journal impact factor: Contingencies and consequences. *Scientometrics, 79*(3), 639–653.

Bar-Ilan, J. (2008). Informetrics at the beginning of the 21st century —A review. *Journal of Informetrics, 2*(1), 1–52.

Campanario, J. M., Carretero, J., Marangon, V., Molina, A., & Ros, G. (2011). Effect of the journal impact factor of the number and document type of citing records: a wide-scale study. *Scientometrics, 87*(1), 75–84.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale,NJ: Erlbaum.

Cronin, B., Snyder, H., & Atkins, H. (1997). Comparative citation rankings of authors in monographic and journal literature: A study of sociology. *Journal of Documentation, 53*(3), 263–273.

Garfield, E. (2009). Five decades of citation indexing. *International Workshop for Scientometrics.* Beijing, China. Retrieved May 30, 2013, from http://garfield.library.upenn.edu/papers/beijingchina2009.html.

Georgas, H., & Cullars, J. (2005). A citation study of the characteristics of the Linguistics Literature. *College & Research Libraries, 66*(6), 496–516.

Gorraiz, J., Purnell, P., & Glänzel, W. (2013). Opportunities and limitations of the book citation index. *Journal of the American Society for Information Science and Technology, 64*(7), 1388–1398.

Hammarfelt, B. (2011). Interdisciplinarity and the intellectual base of literature studies: Citation analysis of highly cited monographs. *Scientometrics, 86*(3), 705–725.

Hicks, D. (2004). The four literatures of social science. In H. F. Moed, W. Glänzel, & U. Schmoch (Eds.), *Handbook of quantitative science and technology research: The use of publication and patent statistics in studies of S&T systems* (pp. 473–496). Netherlands: Kluwer Academic Publishers.

Kousha, K., Thelwall, M., & Rezaie, S. (2011). Assessing the citation impact of books: The role of Google books, Google scholar, and Scopus. *Journal of the American Society for Information Science and Technology, 62*(11), 2147–2164.

Lambiotte, R., & Panzarasa, P. (2009). Communities, knowledge creation, and information diffusion. *Journal of Informetrics, 3*(3), 180–190.

Leydesdorff, L. (2009). How are new citation-based journal indicators adding to the bibliometric toolbox? *Journal of the American Society for Information Science and Technology, 60*(7), 1327–1336.

Leydesdorff, L., & Felt, U. (2012). Edited volumes, monographs and book chapters in the Book Citation Index. *Journal of Scientometric Research, 1*(1), 28–34.

Leydesdorff, L., & Rafols, I. (2009). A global map of science based on the ISI subject categories. *Journal of the American Society for Information Science and Technology, 60*(2), 348–362.

Linmans, A. J. M. (2010). Why with bibliometrics the humanities does not need to be the weakest link. Indicators for research evaluation based on citations, library findings and productivity measures. *Scientometrics, 83*(2), 337–354.

Moed, H. F. (2005). *Citation analysis in research evaluation*. Dordrecht: Springer.

Peritz, B. C. (1981). Citation characteristics in library science: Some further results from a bibliometric survey. *Library Research, 3*(1), 47–65.

Tang, R. (2008). Citation characteristics and intellectual acceptance of scholarly monographs. *College & Research Libraries, 69*(4), 356–369.

Testa, J. (2010).The book selection process for the Book Citation Index in Web of Science. Retrieved June 3, 2013, from http://wokinfo.com/media/pdf/BKCI-SelectionEssay_web.pdf.

Torres-Salinas, D., & Moed, H. F. (2009). Library catalog analysis as a tool in studies of social sciences and humanities: An exploratory study on published book titles in economics. *Journal of Informetrics, 3*(1), 9–26.

Torres-Salinas, D., Robinson-García, N., Jiménez-Contreras, E., & López-Cózar, D. (2012). Towards a 'book publishers citation reports'. First approach using the 'book citation index'. *Revista Española de Documentación Científica, 35*(4), 615–620.

Torres-Salinas, D., Rodríguez-Sánchez, R., Robinson-García, N., Fdez-Valdivia, J., & García, J. A. (2013a). Mapping citation patterns of book chapters in the book citation index. *Journal of Informetrics, 7*(2), 412–424.

Torres-Salinas, D., Robinson-García, N., Campanario, J. M. & Delgado López-Cózar, E. (2013b). Coverage, field specialization and impact of scientific publishers indexed in the 'Book Citation Index'. *Online Information Review*, in press.

White, H., Boell, S. K., Yu, H., Davis, M., Wilson, C. S., & Cole, F. T. H. (2009). Libcitations: A measure for comparative assessment of book publications in the humanities and social sciences. *Journal of the American Society for Information Science and Technology, 60*(6), 1083–1096.

Zuccala, A., & van Leeuwen, T. N. (2011). Book reviews in humanities research evaluations. *Journal of the American Society for Information Science and Technology, 62*(10), 1979–1991.