Journal of Informetrics 3 (2009) 9–26

Contents lists available at ScienceDirect

Journal of Informetrics

journal homepage: www.elsevier.com/locate/joi

Library Catalog Analysis as a tool in studies of social sciences and humanities: An exploratory study of published book titles in Economics

Daniel Torres-Salinas[a], Henk F. Moed[b,*]

[a] *Evaluación de la Ciencia y de la Comunicación Científica, Centro de Investigación Médica Aplicada, Universidad de Navarra, Pamplona, Spain*
[b] *Centre for Science and Technology Studies (CWTS), Leiden University, The Netherlands*

ARTICLE INFO

Article history:
Received 1 August 2008
Accepted 22 October 2008

Keywords:
Library catalogs
Scientific-scholarly books
Social sciences and humanities
Book publishers
Bibliometric indicators
Economics

ABSTRACT

This paper explores the use of Library Catalog Analysis (LCA), defined as the application of bibliometric or informetric techniques to a set of library online catalogs, to describe quantitatively a scientific-scholarly field on the basis of published book titles. It focuses on its value as a tool in studies of Social Sciences and Humanities, especially its cognitive structures, main book publishers and the research performance of its actors. The paper proposes an analogy model between traditional citation analysis of journal articles and Library Catalog Analysis of book titles. It presents the outcomes of an exploratory study of book titles in Economics included in 42 academic library catalogs from 7 countries. It describes the process of data collection and cleaning, and applies a series of indicators and thematic mapping techniques. It illustrates how LCA can be fruitfully used to assess book production and research performance at the level of an individual researcher, a research department, an entire country and a book publisher. It discusses a number of issues that should be addressed in follow-up studies and concludes that LCA of published book titles can be developed into a powerful and useful tool in studies of Social Sciences and Humanities.

1. Introduction

Over the years bibliometrics has proven to be a valid tool in the study of scientific disciplines, especially in Natural and Life Sciences. However the bibliometric methodology, based on counting publications in scientific journals covered by Thomson Scientific's Web of Science and the citations they received has not been effective in many parts of Social Sciences and Humanities. The main reason is that these domains of human scholarship have a series of characteristics and practices that make them different from other research disciplines. Moed (2005) has presented evidence that the coverage of the Web of Science is *moderate* in sociology, political science, anthropology and educational sciences, and particularly in humanities. A principal cause of this moderate coverage is the importance of sources other than international journals, especially books. Moreover, language or national barriers play a much greater role than they do in other domains of science and scholarship. In addition, research activities may be fragmented into distinct schools of thought, each with their own 'paradigms'. He concluded that it would be a great challenge to systematically explore alternative data sources and methodologies for research performance assessments in social sciences and humanities.

* Corresponding author. Tel.: +31 715273940; fax: +31 715273911.
E-mail addresses: torressalinas@gmail.com (D. Torres-Salinas), moed@cwts.leidenuniv.nl (H.F. Moed).

1751-1577/$ – see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.joi.2008.10.002

Van Leeuwen also concluded that the level of reliability and validity of Web of Science based indicators in these areas is low (Van Leeuwen, 2006). Other authors underline the important role of the national orientation of the research (Hicks, 1999) and differences in citation practices (Glanzel & Schoepflin, 1999). Several citation studies confirm the important role of books in the communication of research findings in Social Sciences and Humanities (e.g., Wolfe Thompson, 2002). Books tend to be more frequently cited than papers in scientific journals, although there are significant differences across disciplines, for instance between Philosophy and the Sociology (Lindhom-Romantschuk &Warner, 1996). A review of six studies on various disciplines in Social Science and Humanities conducted by Hicks (1999) concluded that the average numbers of citations received by the books were always higher than those obtained by other types of scientific publications.

The outcomes of these studies reflect a dual situation. On the one hand they stress the importance of books as media of scientific communication, but on the other hand their assessments were mainly based upon an analysis of citations collected from scientific journals. Another fundamental difference between studies of the role of books as compared to that of scientific journals is that the former tend to be based upon collections of books published by individual scholars in universities (Villagrá Rubio, 1992), small research units (Cherchye & Vanden Abeele, 2005) or small disciplines (Lewison, 2001). There are no studies covering the production of books at the macro level.

Bibliometric studies of books are mainly small scale studies because no databases are readily available that allow a systematic, computerised analysis of large sets of books. Although one can find products such as those developed on a commercial basis by *Bowker* and databases developed by national agencies such as the *ISBN database books published in Spain*, these information sources have several severe limitations from a bibliometric point of view. Their primary purpose is to provide information about books for sale; they do not distinguish between scientific and popular books; their interfaces are not ready for downloading records; and they do not give useful information about the affiliations of the publishing authors.

But other bibliographic information sources are available as well. A most important group is formed by virtual library catalogs or OPACS (*Online Public Access Catalogs*). Since the assimilation of the Z39.50 protocol (further explained in Section 3.1) catalogs complying with this protocol have been combined to become large databases with millions of records. The most obvious example of this trend is *WorldCat*, a global online network of library catalogs created by the Online Computer Library Center (OCLC). It is no surprise that research on OPACs and library collections constitutes one of the traditional research fronts in Library Science. These studies have been conducted from multiple perspectives, including that of information retrieval and transactional log analysis. But perhaps the world of bibliometrics is more familiar with the use of citation analysis for the development and optimization of academic library collections (Edward, 1999; Pancheshnikov, 2007; Swygart-Hobaugh, 2004).

Although the latter research topic marks a common point of interest between Library Science (especially the study of library collections) and bibliometrics (especially citation analysis), the exchange of knowledge and research techniques between these two research areas has been rather unidirectional. To the best of our knowledge, bibliometric researchers have never used at a large scale scientific-scholarly library collections in the study of scientific-scholarly development, structure and performance.

2. Conceptualization and objectives

2.1. Conceptualization

Price (1970) conjectured that if a research paper is the expression of one or more persons working at a research front, the papers they produced would be able to provide insight into their research activities and their ties at the time of publication. This premise can be extended to university libraries catalogs in the following way. According to Calhoun (2006) the main mission of library catalogs is to advance the state of knowledge within a library community, and the university library online catalogs reflect a portion of the universe of scholarly information (Calhoun, 2006). While national and international subject bibliographies provide insight into the entire spectrum of what is published, academic library catalogs inform in the first place on the presence of book titles in academic libraries, and therefore primarily reflect the *consumption* in the academic community rather than the *production* of book titles.

The base assumptions underlying the approach proposed in this paper is that the inclusion of a book in academic libraries is an expression of its utility for the academic community, and therefore that an academic library catalog may provide insight into the state of the scientific-scholarly community at a particular time. The larger a collection, the better it approximates a bibliography of a discipline. Combining a number of well selected, specialized library catalogs one may obtain an approximate representation of the production and consumption of information within this discipline. Since the information recorded in the catalogs is adapted to a large number of standards (e.g., *Anglo-American Cataloguing Rules*, *Library of Congress Subject Headings*), one can use the same tools and methods as those applied by bibliometricians to analyse large databases of library catalogs. One could speak of a new type of bibliometric analysis: *Library Catalog Analysis* (LCA).

LCA can be defined as the application of bibliometric techniques to a set of library online catalogs. In this paper LCA is used to describe quantitatively a scientific-scholarly discipline and its actors, on the basis of an analysis of published book titles. Linmans (2008) explored the use of *WorldCat* in an assessment of book titles published by Humanities scholars at Leiden University. The current study further develops and extends this approach. Table 1 establishes an analogy between traditional citation analysis of journal articles and Library Catalog Analysis of book titles.

Table 1
Analogy between citation analysis of journal articles and livrary catalog analysis of book titles.

Citation analysis of journal articles		Library Catalog Analysis of book titles
Main concepts		
Article Author	↔	Book Author/editor
Research article	↔	Book
Publication/citation database	↔	Library catalog
Publication/citation database coverage	↔	Library catalog coverage
Journal Publisher	↔	Book Publisher
Journal's Prestige	↔	Prestige of book publisher or library's institution
Methods and indicators		
Create a set of papers published by a unit	↔	Create a set of book titles published by a unit
Measure the prestige of a journal	↔	Measure the prestige of a publisher or a library s institution
Numbers of citations received by a paper	↔	Numbers of catalog inclusions of a book title
Average number of received citations per paper	↔	Average numbers of catalog inclusions per book title
Compare a unit's citations per paper to an overall or world average	↔	Compare a unit's catalog inclusions per book title to an overall or world average
Geographical spread of authors citing the unit's papers	↔	Geographical spread of catalogs containing the unit's book title
Thematic mapping of publication databases (keywords)	↔	Thematic mapping of library catalogs (subject headings)

A comparison between these two types of bibliometric analyses can be established from two points of view, one conceptual, and one methodological. From a conceptual point of view both have as a unit of analysis publications directed towards the scientific-scholarly community. Citation analysis is normally based on papers published in scientific-scholarly journals, whereas the Library Catalog Analysis explored in this paper focuses on books. At the methodological level many of the techniques employed in citation analysis can be applied to the analysis of book titles contained in OPACS, including the analysis of the prestige of a particular publisher or the geographical spread of a particular work. Perhaps the most attractive element in this analogy is the parallelism between the number of received citations and the number of library catalogs in which books are included. If the main mission of a university library is to advance the state of knowledge for its users, the number of inclusions reflects the usefulness of a particular book. Based on this principle one can build indicators similar to those used traditionally to assess the scientific activity (e.g., Moed, De Bruin, & Van Leeuwen, 1995).

In order to create a database of book titles in a discipline one should select a set of university libraries specialized in that discipline. The final set of book titles collected form their catalogs (Cat. a, b, c and d in Fig. 1) provides a bibliographic representation of the discipline. If additional information is available about the production of books by a particular unit at various levels of aggregation (e.g., an individual scholar, a university department) one can assess their value within the entire database, as shown in Fig. 1.

2.2. Objectives of the study and structure of the paper

The general objective of the study presented in this paper is to explore the potentialities of Library Catalog Analysis, focusing on its value as a tool in studies of social sciences and humanities, especially their cognitive structures, main publishers of book titles, and the performance of the researchers active in these domains of science and scholarship. The aim was to identify the main technical problems and theoretical issues involved, to provide solutions and answers, and finally to make an inventory of issues that need to be further addressed in future research.

From the point of view of data collection and handling a major task is the development of a method for massive downloads of bibliographic records from university library catalogs and the subsequent creation of an off-line database. Hence, the *first*

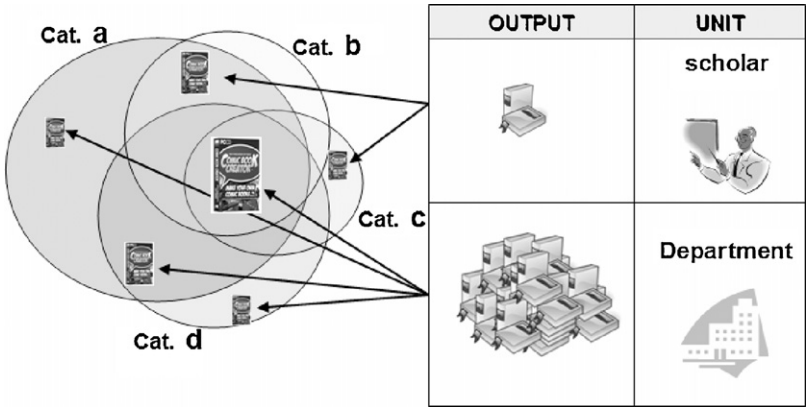

Fig. 1. *Library Catalog Analysis* as a representaion of a research field and as a tool for the analysis of research performance.

objective of the study was to develop such a methodology. This work is described in Section 3.1. It was decided to collect data for one discipline from Social Sciences: Economics. This discipline was selected because both journal articles and books play an important role in its written communication system (Moed, 2005), and it was expected that its core in academic library catalogs could be delimited by means of a simple search string (see Section 3.1 below).

The *second* objective of the study was to develop and apply a series of indicators constructed and interpreted within the framework of the analogy model presented in Table 1. These indicators are defined in Section 3.2.

A *third* objective was to carry out a first bibliometric exploration of the database. Section 4.1 presents a general description of the database, while Section 4.2 analyses the distribution of book titles among catalogs. A series of characteristics of book titles in the database is presented in Section 4.3: their publication language, the number of authors, and the share of authored and edited works. Section 4.4 provides an overview of the research topics covered by the titles in the database, applying a thematic mapping technique based on co-co-word analysis. An analysis of the book publishers is presented in Section 4.5.

Finally, a *fourth* objective was to carry out an exploratory performance assessment study of a particular country, Spain, and the research departments in the field Economics in a particular institution, the University of Navarra (UNAV), a Spanish institution located in Pamplona. The case study of national performance is presented in Section 4.6 and that of institutional performance in Section 4.7.

Chapter 5 gives a discussion of the results, draws main conclusions and highlights a number of important issues that await further research.

3. Materials and methods

3.1. Download from catalogs and creation of an off-line database

The methodology consists of the following steps:

- Selection of tools for information retrieval from university libraries catalogs.
- Selection of university libraries.
- Download of bibliographic records and creation of an off-line database.
- Standardization and cleaning of the database.

A well known standard in the library environment is *ISO 23950: Information Retrieval (Z39.50): Application Service Definition and Protocol Specification* and *ANSI/NISO Z39.50* or just *Z39.50*. Its main objective is to specify a client–server protocol for searching and retrieval of information from remote databases, in this case library systems. It enables one to launch searches from a Z39.50 client to a Z39.50 server. A Z39.50 client-allows one to search in a Z39.50 library server through a variety of attributes including the author name, publication title, ISBN, publication date and subject heading. The protocol ensures that these two different systems can exchange information (Evans, 2001a, 2001b). The records retrieved in a search can also be downloaded to a client in various MARC formats (MARC 21, UNIMARC, IBERMAC, the Library of Congress standards for the representation and communication of bibliographic information in machine-readable form). Therefore, the final set of retrieved records contains highly structured information. The study presented in this paper used *Bookwhere Academic 6.0*. This software supports searches under subject headings, searches in several libraries at the same time and it exports records in various formats (*xml*, *txt*). Another information source used in the study is the Copac Library Catalogue, a combined catalog that includes the major university libraries in the United Kingdom and Ireland, including the British Library.

For the selection of university libraries the following criteria were applied:

- *The Z39.50 protocol should be implemented*. Not all university libraries have implemented this protocol in their OPACs. Z39.50 compliance is unevenly distributed across countries. It is frequently used in the Anglo-Saxon world (United States, United Kingdom, Canada).
- *The library should be a part of a university with a leading position in Economics*. Only Universities were selected with a significant article production and citation impact in this field according to the listings in Thomson Scientific's *Essential Science Indicators*.
- *A Spanish Emphasis*. Since the case studies of research performance assessment relate to Spain and the Spanish University of Navarra it was considered appropriate to extend the sample of Spanish university libraries.

Applying these three criteria, the total number of libraries selected was 42. In these catalogs a search was carried out for the string *ECONOM** in the subject headings of book titles. This truncated search retrieved records indexed as *Economist*, *Economy* or *Econometrics,* but also as *Economic Conditions* or *Medical Economics.* The chronological period of searches was limited to the years 1995–2005. Downloaded records were exported as text files and integrated into a relational database built in MS ACCESS. The main table of the database contains bibliographic descriptions of book titles and the names of the university libraries in which they are included. Records were downloaded during the time period December 2007–January 2008.

In a next step the database was cleaned. Only those records were kept that had a ISBN, i.e., a code given by the various ISBN Agencies to books for sale. The ISBN was used as a unique identifier of a book title, taking into account the different

ISBN numbers that books may have according to their format (*paperback*, *hardback*, *electronic copy*, etc.). Two new data fields were created on the basis of information from existing fields: the number of authors, and a field indicating whether or not the book was an edited work. The following data fields were normalized: publisher, publication language and publication year.

Finally, a file was included containing the titles of books and book chapters published by researchers in the departments of Economics at the University of Navarra during the period 1999–2005. This Spanish university hosts three departments that are active in Economics (Economía, Métodos Cuantitativos and Empresa Informativa) and the IESE Business School, an international institute with sites in Madrid, Barcelona and the United States.

3.2. Indicators

Within the framework of the analogy model presented in Table 1 the following indicators were constructed.

3.2.1. Number of titles (NT)

The total number of unique book titles for a particular aggregate. Typical examples are the number of titles published by a publisher or by researchers from a particular institute. Even if a catalog contains more than one copy of a particular book title, this title is counted as one in all analyses presented in this paper.

3.2.2. Catalog inclusions (CI)

The total number of catalog inclusions for a given set of book title(s). It indicates the dissemination of a (given set of) book title(s) among university libraries. If n_j indicates the number of catalogs in which book title j is included, CI is defined as follows:

$$\mathrm{CI} = \sum_{j=1}^{\mathrm{NT}} n_j.$$

3.2.3. Catalog inclusion rate (CIR)

This indicator calculates for a given set of book titles the average number of catalog inclusions per book title. It is defined as the ratio of *the total number of catalog inclusions* and the *number of titles in the set.*

$$\mathrm{CIR} = \left(\frac{\mathrm{CI}}{\mathrm{NT}}\right)$$

3.2.4. Relative Catalog Inclusion Rate (RCIR)

This measure facilitates the comparison of *Catalog Inclusion Rates* across different aggregates correcting for differences in the number of published book titles. It is defined as the ratio of CIR of the aggregate to be assessed and CIR of the aggregate that serves as a benchmark in the assessment. A special case is the calculation of a RCIR in which CIR of an institute under assessment is divided by the CIR calculated for the total database. The latter parameter can be interpreted as a 'word average'. A value above one indicates that an institution's Catalog Inclusion Rate is above world (or total database) average. This indicator can be conceived as the analogo of relative or normalized citation impact indicators often used in citation analysis. It is expressed as

$$\mathrm{RCIR} = \left(\frac{\mathrm{CIR_a}}{\mathrm{CIR_b}}\right)$$

where $\mathrm{CIR_a}$ = *Catalog Inclusion Rate* of the aggregate to be assessed; $\mathrm{CIR_b}$ = *Catalog Inclusion Rate* of the 'benchmark' aggregate used for comparison.

3.2.5. Diffusion rate (DR)

The percentage of catalog inclusions of book titles produced by a given aggregate relative to the total number of possible catalog inclusions. The number of possible inclusions is equal to the product of the number of titles in the set and the number of catalogs included into the analysis. DR values range between 0 and 1. A value of 1 indicates that each title analysed is present in all library catalogs taken into account. This indicator can be applied to large sets of books or to individual titles, and is defined as:

$$\mathrm{DR} = \left(\frac{\mathrm{CI_s}}{\mathrm{CI_m}}\right)$$

where $\mathrm{CI_s}$ = *Catalog Inclusions* for a given set of titles; $\mathrm{CI_m}$ = maximum number of possible inclusions that the set of titles can reach.

The exploratory study presented in this paper also applied other bibliometric and data-analytical techniques. *Multidimensional Scaling* (MDS) was applied to analyse the similarity between libraries based on the number of titles a pair of libraries

Table 2
Overall statistics of the database of book titles in Economy.

Overall statistics	
Number of countries	7
Number of host Universities	42
Number of unique book titles	121,147
Number of catalog inclusions	417,033
Overall Catalog Inclusion Rate	3.4

has in common. A second technique is an extension of *co-word analysis* (Callon, Courtial, & Laville, 1991; Callon, Courtial, Turner, & Bauin, 1983) developed by Bailón-Moreno (2003) in the software *Copalred 1.0*. It was used to uncover the main themes of books as reflected in the subject headings assigned to them in the various library catalogs. *Social Network Analysis* is a third technique applied in this paper. It is used to graphically display and analyse the geographical spread of books published by the University of Navarra.

4. Results

4.1. General database description

Table 2 presents overall statistics of the database created in the study. The total number of university libraries catalogs included in the study was 42. Table 3 shows that they are located in 7 countries. The country with the largest number of catalogs is Spain with 12 followed by United States and the United Kingdom with 10.

The total number of titles collected was 121,147, accounting for a total of 417,033 inclusions. Hence, the *Catalog Inclusion Rate* for the total set of titles is 3.4 catalogs per unique book title. Table 3 shows that the USA is the country that has the largest collection with 71,731 titles. In a ranking of libraries according to the size of their collections, eight libraries contain over 20,000 titles, including six US libraries led by *University of Michigan Library* and *Yale University Library*. In Canada *The University of British Columbia Library* and in the United Kingdom the *Library of the London School of Economics and Political Science* exceed this threshold. Spanish libraries tend to have relatively small collections. Appendix A provides a list of all 42 libraries included.

A MDS map presented in Fig. 2 shows an image of the 42 libraries. The libraries are located in a two-dimensional plane according to their similarity among their collections. Similarity between a pair of catalogs is defined using Salton's Index, i.e., the number of titles they have in common, divided by the square root of the product of the number of titles in each catalog. One can distinguish between the following clusters. At the upper side of the map one finds in cluster I the Spanish libraries; those from universities in Madrid are located on the right side. They are all remote from the Anglo-Saxon libraries. The central cluster II is formed by the libraries with large collections with more than 20,000 titles, most of them from the United States. Cluster III in lower part of the graph consists of a group of medium-sized British libraries. Other libraries are more dispersed and are not linked to a specific geographical area.

Table 4 shows the global production of book titles in Economics during the time period 1995–2005. The number of titles published annually exceeded 10,000 in all years; the annual average amounts to 11,415. Between 1995 and 2000 there was an increase in the annual number of titles published, followed by a decline as from 2001. The *Catalog Inclusion Rate* exceeded 3.10 in all years. The highest rate was obtained in the beginning of the period (3.6 inclusions per unique book title).

No information was available on the date at which titles were included in the catalogs. Hence, it was not possible to estimate the time delay with which published books were included in the catalogs. Therefore, it could not be established whether the decline in number of titles in recent years reflects a genuine decline in the annual number of titles purchased by the 42 libraries analysed in this study, or whether it is mainly caused by the time delay between publication date and inclusion date.

4.2. Distribution of book titles among catalogs

Table 5 analyses the sets of catalogs from all countries combined, and those from the USA and from European countries. Each analysis takes into account book titles with at least one catalog inclusion in a particular set. Table 5 shows for all

Table 3
Distribution of analysed library catalogs by their country of origin.

Country of library catalog	No. of catalogs	No. of titles	Catalog inclusions
Australia	4	25,132	34,500
Canada	4	34,184	52,594
Ireland	1	3,395	3,395
Spain	12	17,164	26,254
Sweden	1	4,074	4,074
United Kingdom	10	44,490	85,696
United States	10	71,731	210,529

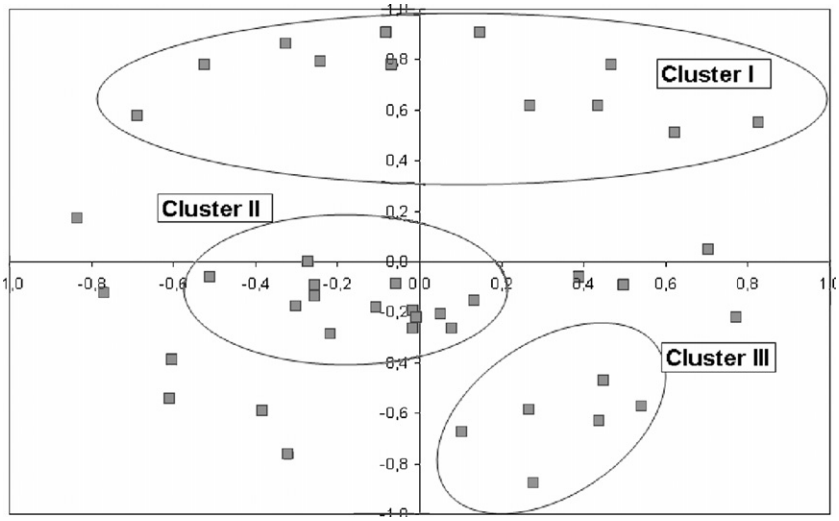

Fig. 2. Multidimensional Scaling Map based on the degree of overlap among.

countries combined that 45 percent of the titles appear in one catalog only. There are notable differences between US and European libraries. For US libraries 40 percent of titles are included in only one catalog, whereas for European libraries it is 57 percent. Generally, the degree of concentration of book titles among libraries is larger in the US libraries than it is in their European counterparts. This outcome probably reflects a greater cultural and linguistic homogeneity in the US as compared to Europe.

This difference is also evident in the tail of the distribution. In the total collection the maximum number of inclusions of a title is 33 out of 42; only one book reaches this value (see Table 6 for the bibliographic details of this book). In the United States 1.7 percent of titles appear in each of the 10 libraries analysed in the study. In Europe the maximum number of inclusions is 17 out of 24, and there is only one book that reaches this maximum.

Fig. 3 shows a graphical display of these distributions in a double-logarithmic plot. The curve related to the total set of catalogs reveals a configuration of two approximately straight lines with different slopes, representing two power-law distributions with exponents −1.48 and −9.18, respectively. Fig. 3 shows that the left part of the curve is largely determined by the US libraries, and the right part by their European counterparts.

Table 6 presents the book titles in Economics with the largest number of catalog inclusions in the set of 42 analysed in this study. *Numerical Methods in Economics* by Judd Kenneth and published by MIT Press in 1998 ranks on top. It is catalogued included in as many as 33 libraries and its *Diffusion Rate* is 0.78. Second and third are *In Defence of Globalization* by J.N. Bhagwatti (Oxford University Press, 2004) and *kicking away the Ladder* by H-J. Chang (Anthem, 2002), both with 30 inclusions. The books in Table 6 tend to have rather general titles. Some provide an introduction to specific sub-disciplines or topics that may play an important role in teaching. Others focus more on research such as handbooks and provide overviews or introductions of specific topics or research methods (e.g., numerical methods, computational Economics). Many titles relate to politically relevant issues including globalization, ethics and work and leisure.

4.3. Publication language, number of authors per title, and number of edited works

Books have been published in a total of 89 different languages, but Table 7 shows that the ten most important languages account for more than 90 percent of all titles. The dominant position of English at least partly reflects the Anglo-Saxon bias

Table 4
Trend in annual number of book titles during 1995–2005.

Publication year	No. of titles	Catalog inclusions	Catalog Inclusion Rate
1995	11,170	40,296	3.6
1996	11,494	41,784	3.6
1997	11,269	39,011	3.5
1998	11,410	39,376	3.4
1999	11,737	39,016	3.3
2000	12,728	39,970	3.1
2001	12,558	39,077	3.1
2002	11,677	37,101	3.2
2003	10,695	35,829	3.3
2004	10,412	33,701	3.2
2005	10,414	32,539	3.1

Table 5
Distribution of book titles among catalogs for US, European and all libraries combined.

All countries combined (no. of catalogs = 42)				United States (no. of catalogs = 10)			European Countries (no. of catalogs = 24)		
No. of catalog inclusions	No. of titles	%Titles	Cumm% titles	No. of titles	%Titles	Cumm% titles	No. of titles	%Titles	Cumm% titles
1	54,978	45	45	28,872	40	40	33,531	57	57
2	19,424	16	61	12,851	18	58	10,703	18	75
3	11,152	9.2	71	8,121	11	70	5,941	10	85
4	7,547	6.2	77	6,193	8.6	78	3,515	6.0	91
5	5,452	4.5	81	4,758	6.6	85	2,171	3.7	95
6	3,915	3.2	85	3,143	4.4	89	1,324	2.2	97
7	3,288	2.7	87	2,558	3.6	93	768	1.3	98
8	2,493	2.1	89	2,034	2.8	96	461	0.8	99
9	2,238	1.8	91	2,016	2.8	98	249	0.4	100
10	1,757	1.5	93	1,185	1.7	100	122	0.2	100
11	1,585	1.3	94				65	0.1	100
12	1,401	1.2	95				21	0.0	100
13	1,240	1.0	96				9	0.0	100
14	1,097	0.9	97				8	0.0	100
15	907	0.7	98				3	0.0	100
16	743	0.6	98				1	0.0	100
17	544	0.4	99				1	0.0	100
18	397	0.3	99						
19	330	0.3	100						
20	245	0.2	100						
21	145	0.1	100						
22	93	0.1	100						
23	69	0.1	100						
24	51	0.0	100						
25	30	0.0	100						
26	12	0.0	100						
27	6	0.0	100						
28	5	0.0	100						
29	33	1	0.0						
30	2	0.0	100						
33	1	0.0	100						

in the selection of libraries analysed in the study. In libraries located in Australia, Ireland and the United Kingdom more than 90% of book titles is written in English. However, USA shows a different picture. In the ten libraries from this country included in the study, 68 percent of titles are in English, followed by 7.2% in Spanish and 4.5 percent in Chinese. In the case of Spanish libraries the situation is different. A majority of 52 percent of their titles is written in Spanish, but noteworthy is a high share of 38 percent of books in English. This outcome reflects the importance of English in the field of Economics, but also the openness of Spanish academic institutions towards other cultural and language domains.

During the time period 1995–2005 the average number of authors per book title amounts to 1.63. This value remained almost constant over time during the time period considered. Therefore the number of authors is not growing in books as in the case in co-authorship in journal articles. Table 8 reveals the importance of single authored books in Economics: 62 percent of the titles were written by one author only, 20 percent by two, and 13 percent by three authors.

In 90 percent of the bibliographic records in the database there is an indication whether of not a book is an *edited* work. An editor is responsible for selecting the authors of chapters and ensures the quality of the content. The results indicate that

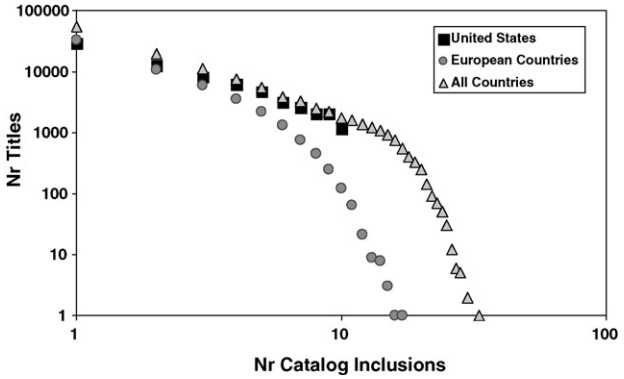

Fig. 3. Distribution of book titles among catalogs.

Table 6
The 25 book titles with the highest number of inclusions in the set of 42 academic libraries.

Author	Title	Publisher	Edited	Date	No. of catalog inclusions	Diffusion rate
Judd, Kenneth L	Numerical methods in Economics	MIT Press	No	1998	33	0.786
Bhagwati, Jagdish N	hi defense of globalization	Oxford University Press	No	2004	30	0.714
Chang, Ha-Joon	Kicking away the ladder	Anthem	No	2002	30	0.714
Sen, Amartya Kumar; Basu, Kaushik; Pattanaik, Prasanta K; Suzumura, Kotaro	Choice, welfare, and development: A Festschrift in Honour of Amartya K. Sen	Oxford University Press	Yes	1995	28	0.667
Fujita, Masahisa; Krugman, Paul R; Venables, Anthony	The spatial economy	MIT Press	No	2000	28	0.667
Begg, Iain; Henry, S G B	Applied Economics and public policy	Cambridge University Press	Yes	1998	28	0.667
Maddison, Angus	The World Economy: A Millennial Perspective	Development Centre of the OBCD	No	2001	28	0.667
Helpman, Elhanan	The mystery of economic growth	Harvard University Press	No	2004	28	0.667
Gershuny, lonathan	Changing times: work and leisure in postindustrial society	Oxford University Press	No	2000	27	0.643
Brennan, Geoffrey; Pettit, Philip	The economy of esteem: An Essay on Civil and Political Society	Oxford University Press	No	2004	27	0.643
Little, Ian Malcolm David	Ethics, Economics, and politics	Oxford University Press	No	2002	27	0.643
Alesina, Alberto; Roubini, Nouriel; Cohen, Gerald D	Political cycles and the macroeconomy	MIT Press	No	1997	27	0.643
Epstein, Gerald A; Gintis, Herbert	Macroeconomic policy after the conservative era	Cambridge University Press	Yes	1995	27	0.643
lames, Harold	The end of globalization	Harvard University Press	No	2001	27	0.643
Bonney, Richard	Economic systems and state finance	Clarendon Press	Yes	1995	26	0.619
Clark, Gordon L;;Feldman, Maryann P;Gertler, Meric S	The Oxford handbook of economic geography	Oxford University Press	Yes	2000	26	0.619
Chang, Ha-loon; Rowthorn, Bob; World Institute for Development Economics Research	The role of the state in economic change	Clarendon Press-Oxford University Press	Yes	1995	26	0.619
Kay, IA	The business of Economics	Oxford University Press	No	1996	26	0.619
Obstfeld, Maurice; Rogoff, Kenneth S	Foundations of international macroeconomics	MIT Press	No	1996	26	0.619
Amman, Hans M; Kendrick, David A; Rust, lohn	Handbook of computational Economics	Elsevier	Yes	1996	26	0.619
McCormick, Michael	The origins of the European economy	Cambridge University Press	No	2001	26	0.619
Castells, Manuel	End of Millenium: The Information Age	Blackwell Publishers	No	2000	26	0.619
Lucas, Robert E	Lectures on economic growth	Harvard University Press	No	2002	26	0.619
Putnam, Hilary	The collapse of the fact/value dichotomy and other essays	Harvard University Press	No	2002	26	0.619
Nelson, Richard R	The sources of economic growth	Harvard University Press	No	1996	26	0.619

15 percent of book titles are edited works. Analyzing this percentage as a function of the number of catalog inclusions it was found that edited words are overrepresented in the set of titles with a large number of catalog inclusions. For instance, among the 240 titles with 20 or more inclusions 36 percent is an edited work, whereas in the subset of books appearing in one catalog only this percentage drops to 9 percent. Fig. 4 reveals this pattern graphically.

4.4. Thematic mapping using co-word analysis

In the same way as one creates thematic maps based on articles in scientific journals and patents, it is possible to use existing mapping techniques to reveal the topics of book titles in Economy. Fig. 5 shows a typical example. It applies co-word analysis to the subject headings of book titles in libraries located in USA and UK.

Table 7
Publication languages of book titles in libraries located in the various countries.

All 42 libraries			%Titles in the country where libraries are located						
Language	No. of titles	%Titles	Australia	Canada	Ireland	Spain	Sweden	UK	USA
English	75211	61	93	86.5	*922*	38.1	86.5	89	67.8
Spanish	13348	10.8	0.9	1.6	2.8	52.4	2.4	1.7	7.2
French	6209	5.0	0.9	3.8	2.8	4.0	0.7	2.6	3.5
Chinese	5858	4.8	0.7	3.4	0.0	0.0	0.0	0.1	4.5
German	4579	3.7	0.3	1.0	0.5	0.9	0.4	1.5	3.9
Russian	3104	2.5	0.1	0.4	0.0	0.0	0.1	1.4	2.7
Italian	2344	1.9	0.1	0.3	0.3	1.1	0.0	0.9	1.7
Japanese	2191	1.8	0.8	1.0	0.1	0.0	0.1	0.0	1.7
Indonesian	1167	1.0	1.2	0.8	0.0	0.0	0.0	0.0	1.3
Portuguese	1142	0.9	0.0	0.0	0.0	0.3	0.1	0.2	1.0

Table 8
The number of authors of book titles.

No. of authors	No. of titles	%Titles
1	74,952	62
2	24,853	21
3	15,424	13
4	4,208	3.5
5	1,013	0.84
6	507	0.42
7	88	0.07
8	62	0.05
9	16	0.01
10	18	0.01
11	6	0.00
12	7	0.01
13	1	0.00
14	1	0.00

In the strategic diagram presented in Fig. 5 the vertical axis measures the density—i.e., the strength of the internal links within a cluster represented by a theme—, and the horizontal vertical axis the centrality—i.e. the strength of the links between the theme and other themes in the map. Hence, the themes with the highest internal coherence and closest relationship to other themes appear in the first quadrant (the upper right part of the graph). In this quadrant the following topics can be found: *Medical Economics*, *History*, *Agriculture*, *Urban Economics*, *International Relations* and *Economics Environmental Policy*. As a further illustration of the co-word technique Fig. 6 displays sub-topics and their interrelationships in one particular main theme: *International Economic Relationships*.

4.5. Publisher performance

A total of about 22,000 publishers have published at least one book title catalogued in at least one of the 42 catalogs under a Subject Heading containing the string *Econom**. Fig. 6 presents the distribution of the number of book titles amongst publishers. 61‘percent of publishers have published only one book; these publishers are not specializing in the field Economics.

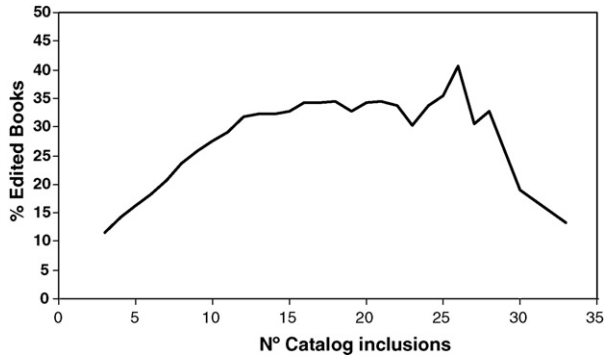

Fig. 4. Percentage of edited books as a function on the number of catalog inclusions.

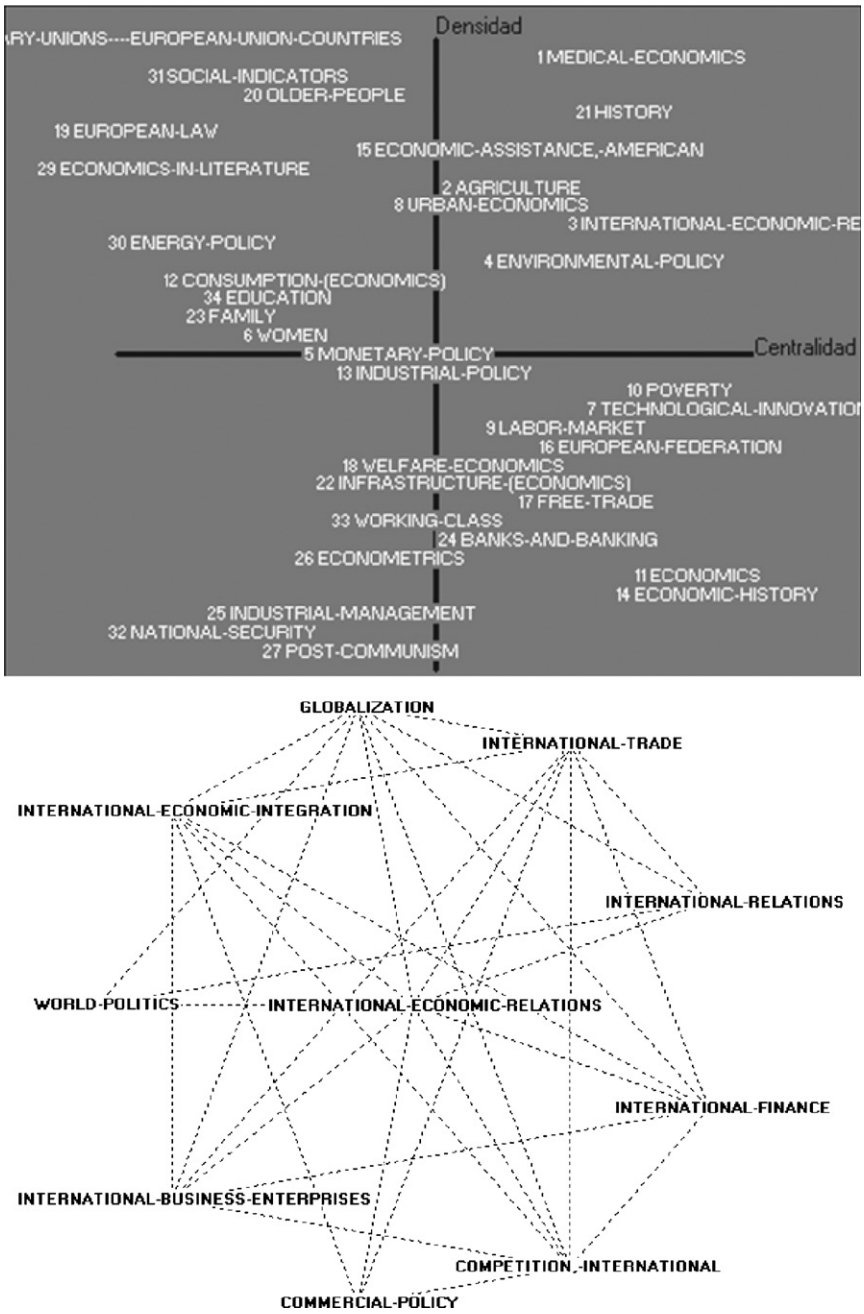

Fig. 5. Thematic map of the subject headings of book titles in US and UK libraries: strategic diagram and thematic example.

Only 133 publishers published more than 10 titles per year. The distribution is similar to other bibliometrics distributions such as that relating to the article productivity of individual authors.

Table 9 shows the names of the 25 most productive publishers. All of them have published more than 400 titles during the time period 1995–2005. It clearly demonstrates that there are two types of publishers: commercial publishers and those associated with national or international organizations or agencies. Within the set of commercial publishers the most important ones are *Routledge*, *Edward Elgar* and *Oxford University Press*. Among the publishing organizations the *World Bank* leads, but the list also includes national organizations or agencies such as *Her Majesty's Stationary Office* in the United Kingdom. The American publisher *St. Martin's Press* obtains the highest Catalog *Inclusion Rate* with 7.7 catalogs per book title, followed by *Palgrave* with 7.5 and *Cambridge University Pres* with 7.2. It must be noted that the standardization of publishers is not complete. Rather recent mergers, such as that of Springer and Kluwer were *not* taken into account.

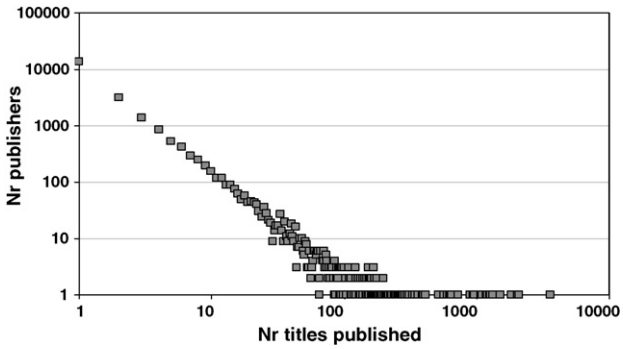

Fig. 6. Distribution of number of book titles amongst publishers.

Publishers' *Diffusion Rate* was calculated for the aggregate of all libraries but also per geographical area. A high DR value tends to go along with a high value of CIR. Calculation of this measure by geographical area allows one to identify the area of influence of the various publishers. Thus the publishers from government offices tend to have only influence in their national territory. *United States GPO* shows an overall DR of 0.16, but one can see that this rate is much higher in the USA than it is in Europe (0.67 versus 0.0). Another example, in the set of commercial publishers, is *St. Martin Press* whose importance is quite remarkable in the United States with a DR of 0.45, but much lower in Europe. The *Diffusion Rate* tends to be smaller in European countries than in other countries due to the effect of the 12 Spanish libraries included in the study, in which the number of titles in English is lower and international publishers have a smaller presence. It reflects the situation in Europe which is much more heterogeneous in terms of publishers and where national publishers play an important role.

Fig. 7 shows for publishers publishing at least 10 titles per year the number of titles per year (measuring productivity) and their *Relative Catalog Inclusion Rate* (measuring visibility). It clearly shows a group of prominent publishers each with more than 150 titles per year and a RCIR exceeding 1.5. A second group contains 18 publishers publishing between 50 and 150 titles, the major part of which has a RCIR value above the database or 'world' average. A third group includes publishers with fewer than 50 titles and a variable RCIR.

Table 9
The 25 most productive publishers in Economics.

Publisher name	Total no. of titles	No. of titles per year	No. of Catalog inclusions	Catalog Inclusion Rate	Relative Catalog Inclusion Rate	Diffusion Rates		
						All countries DR	European Countries DR	United States DR
Routledge	3589	326	19,985	5.6	1.6	0.13	0.10	0.24
Edwad Elgar	2102	191	13,922	6.6	1.9	0.16	0.13	0.22
Oxford University Press	2092	190	14,576	7.0	2.0	0.17	0.14	0.29
Cambridge University Press	2040	185	14,710	7.2	2.1	0.17	0.15	0.29
Macmillan Press	1827	166	10,054	5.5	1.6	0.13	0.09	0.25
World Bank	1509	137	8,612	5.7	1.7	0.14	0.07	0.30
Official Publications of the EC	1344	122	6,528	4.9	1.4	0.12	0.12	0.17
United States GPO	1281	116	8,613	6.7	2.0	0.16	0.00	0.67
OECD	1260	115	6,308	5.0	1.5	0.12	0.09	0.19
Her Majesty's Stationery Office	1209	110	2,315	1.9	0.6	0.05	0.08	0.02
Palgrave	1188	108	8,905	7.5	2.2	0.18	0.11	0.37
United Nations	1109	101	4,993	4.5	1.3	0.11	0.05	0.25
Springer	1078	98	3,995	3.7	1.1	0.09	0.06	0.12
Ashgate	966	88	5,992	6.2	1.8	0.15	0.10	0.24
Kluwer	944	86	4,133	4.4	1.3	0.10	0.08	0.17
St. Martin's Press	894	81	6,904	7.7	2.2	0.18	0.09	0.45
Harmatan	727	66	1,506	2.1	0.6	0.05	0.02	0.15
International Monetary Fund	646	59	2,856	4.4	1.3	0.11	0.09	0.15
McGrawHill	645	59	2,111	3.3	1.0	0.08	0.10	0.07
Wiley	633	58	2,380	3.8	1.1	0.09	0.08	0.13
MIT Press	632	57	3,984	6.3	1.8	0.15	0.13	0.22
Sage	606	55	2,995	4.9	1.4	0.12	0.09	0.21
Blackwell	574	52	3,318	5.8	1.7	0.14	0.13	0.19
Prentice Hall	518	47	2,060	4.0	1.2	0.10	0.11	0.10
Earthscan	416	38	2,352	5.7	1.6	0.14	0.14	0.19

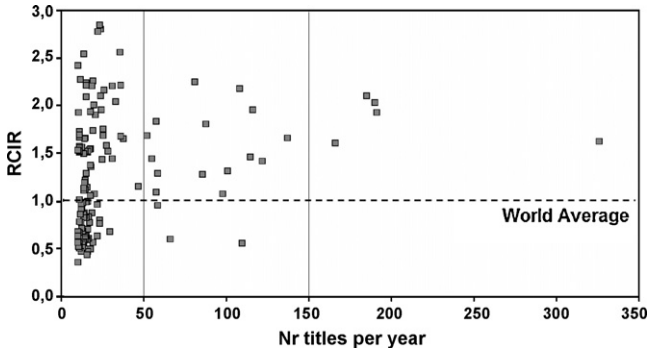

Fig. 7. Number of titles published per year and Relative Catalog Inclusion Rate (*RCIR*) for major publishers.

Table 10
Indicators of Spanish book titles in Economics.

	No. of titles	%Titles	No. of Catalog Inclusions	Catalog Inclusion Rate	Relative Catalog Inclusion Rate	Diffusion Rate
World	6704	100	1,493	2.2	0.6	0.05
United States	740	110	1,494	*2.0*	0.6	002
Spain	6704	100	13,016	1.9	1.2	0.16
United Kingdom	197	2.9	225	1.1	0.5	0.00
North America	763	11.3	1,603	2.1	0.6	101
European countries	6704	100	13,303	1.9	0.9	0.09

4.6. National performance

This section analyses the book titles produced by Spain in Economics. In this section a Spanish book was defined as a book of which the ISBN starts with the digits 84, the country code assigned to Spain. Applying this criterion, Table 10 shows that the total number of books produced by Spain during 1995–2005 amounts to 6704, a number that represents 5.5 percent of the database total. Of these 6704 Spanish books, 11 percent is included in US libraries, and only 2.9 percent in libraries located in the UK.

The worldwide Catalog Inclusion Rate for Spanish book titles is 2.2 catalogs per title, which is below the reference value of 3.4 calculated for the entire database. In fact, the Relative Catalog Inclusion Rate of Spanish book title is below 1.0 in each geographical area outside Spain. Diffusion Rates are also rather low, except within Spain itself.

Table 11 shows the time evolution of the various indicators of Spanish book title production. The average number of titles per year is 653. During the time period 1995–2005 the annual number of titles follows the same pattern as that for the total collection presented in Table 4: a slight increase up until the year 2000 followed by a decline as from the year 2001. The *Inclusion Rate* reached a maximum value of 2.5 in the year 2002.

Table 12 presents a list of Spanish book titles with more than 8 inclusions. The maximum number of inclusions is 11 and relates to *Un siglo de España. La economia* by L.J. Garcia-Delgado and J.C. Jiménez-Jiménez. It is present in 32 percent of the libraries. Table 12 also presents Diffusion Rates within the various countries. All titles have Diffusion Rates within the USA of at least 0.6. In most cases these rates are higher than those calculated for diffusion within Spain. The book titles listed in Table 12 relate to Spanish America, or to historical topics.

Table 11
Time evolution in indicators of Spanish book production.

Year	No. of titles	World share (%)	No. of catalog inclusions	Catalog Inclusion Rate	Diffusion Rate
1995	611	5.4	1433	2.3	0.05
1996	656	5.7	1648	2.5	0.06
1997	637	5.6	1532	2.4	0.05
1998	679	5.9	1644	2.4	0.05
1999	719	6.1	1729	2.4	0.05
2000	726	5.7	1778	2.4	0.05
2001	696	5.5	1695	2.4	0.05
2002	674	5.7	1723	2.5	0.06
2003	659	6.1	1578	2.3	0.05
2004	591	5.6	1259	2.1	0.05
2005	535	5.1	1116	2.0	0.05

Table 12
Spanish book titles with more than 8 inclusions.

Bibliographic description	No. of Catalog Inclusions	Diffusion Rates		
		All countries DR	United States DR	Spain DR
Luis Jose Garcia Delgado; Juan Carlos Jimenez Jimenez. *Un Siglo De Espaha. ha economia.* Marcial Pons. 1999	11	0.32	0.60	0.62
Santamaria Garcia. Antonio; Malamud. Carlos. Sin aziicar no hay pals la industria azucarera y la economia cubana (1919–1939). Universidad de Sevilla. CSIC. 2001	10	0.29	0.60	0.50
Piqueras Arenas. Jose A. Cuba, emporio y colonia la disputa de un mercado interferido (1878–1895) Fondo de Cultura Economica de España. 2003	9	0.26	0.60	0.37
Orti Gost. Pere. Renda ifiscalitat en una ciutat medieval. CSIC. 2000	8	0.23	0.70	0.12
Gutierrez Escudero. Antonio.; Martinez Ortega. Ana Isabel. *Ciencia. economia y politico en Hispanoamerica colonial.* CSIC. 2000	8	0.23	0.60	0.25
Maluquer de Motes Bernet. Jordi. Espana en la crisis de 1898 de la gran depresion a la modernizacion economica del siglo XX. Peninsula. 1999	8	0.23	0.60	0.25
Vigo Gutierrez. Abelardo. *Cambistas. mercaderes y banqueros en el Siglo de Oro espanol* Biblioteca de Autores Cristianos. 1997	8	0.23	0.60	0.25
Villarias Robles. Juan Jose R. *El sistema economico del imperio inca.* CSIC. 1998	8	0.23	0.70	0.12

4.7. Institutional performance

The departments of Economics at the University of Navarra (UNAV) have published during 1995–2005 a total of 211 works, 48 percent of these were monographs or edited works, and 52 percent book chapters. The UNAV book file was de-duplicated. If there were two or more contributions to an edited work, this work was counted only once. 66 percent of publications carried a Spanish ISBN. In this first exploratory analysis no distinction was made between monographs, edited works and book chapters. All types were given equal weights. Table 13 presents the results for the total production of UNAV. 44 titles were found in at least one library, 29 are in present in at least one Spanish library and 11 in at least one library located in the USA.

The total number of catalog inclusions is 142 and the Catalog *Inclusion Rate* is 3.2, which is slightly lower than this rate calculated for the entire database. In fact, the Relative Catalog Inclusion Rate amounts to 0.9. The Diffusion Rate within USA is somewhat higher than that within Spain (2.8 versus 2.5) but Spain has the highest *RCIR* (1.6 versus 0.9).

Comparing the outcomes for the University of Navarra with those for Spain (see Fig. 8) it can be concluded that the university's *Relative Inclusion Rate* and the *Diffusion Rate* are always above those for Spain, except the indicator measuring the *Diffusion Rate* within Spain. The largest differences occur in the inclusion and diffusion within North American institutions, where UNAV book titles have a greater visibility.

The distribution of the number of inclusions across catalogs is similar to that obtained for the various geographical aggregates presented in Table 5. Table 14 shows that in the case of the University of Navarra 54 percent of the titles are just in one catalog. The largest number of inclusions a title achieved amounts to 23. From the point of view of visibility in libraries this is the most important title among those published by researchers at UNAV.

Another way to analyse the geographical spread of book titles is through the use of social network analysis. A typical example is shown in Fig. 9. In this figure a small, light gray node represents a single title. Its diameter indicates the number of catalogs in which it is included. Countries are represented by bigger, dark grey circles. An inclusion of a book title in a library

Table 13
Indicators of the production of book titles in Economics from the University of Navarra.

	No. of titles	%Titles	No. of Catalog Inclusions	Catalog Inclusior Rate	Relative Catalog Inclusion Rate
Entire database	44	100	142	3.2	0.9
United States	11	25	31	2.8	0.9
Spain	29	66	72	2.5	1.6
United Kingdom	12	26	20	1.7	0.9
North America	12	27	43	3.6	1.1
European countrk	40	91	98	2.4	1.2

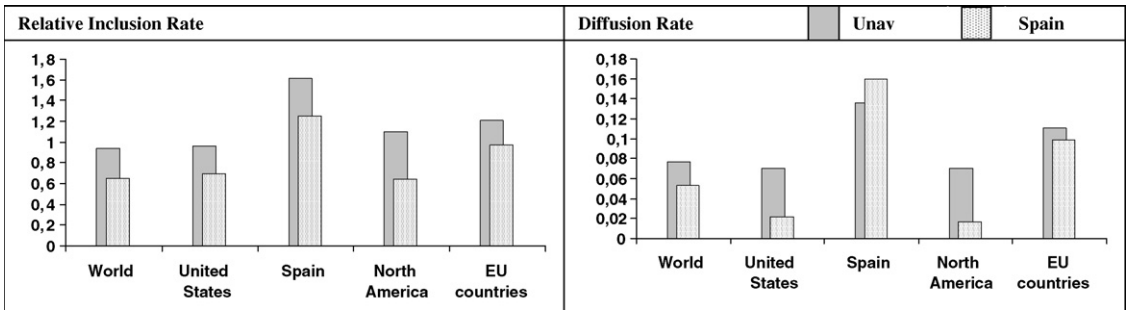

Fig. 8. Comparison of *Relative Catalog Inclusion Rates* and *Diffusion Rates* between the University on Navarra (Unav) and Spain.

Table 14
Distribution of book titles from University of Navarra among catalogs.

No. of Catalog Inclusions	No. of titles	%Titles	Cumm% titles
1	24	54.3	54
2	5	11.2	65
3	3	6.8	72
4	3	6.8	79
6	4	9.1	88
8	2	4.5	93
10	1	2.3	95
14	1	2.3	97
23	1	2.3	100

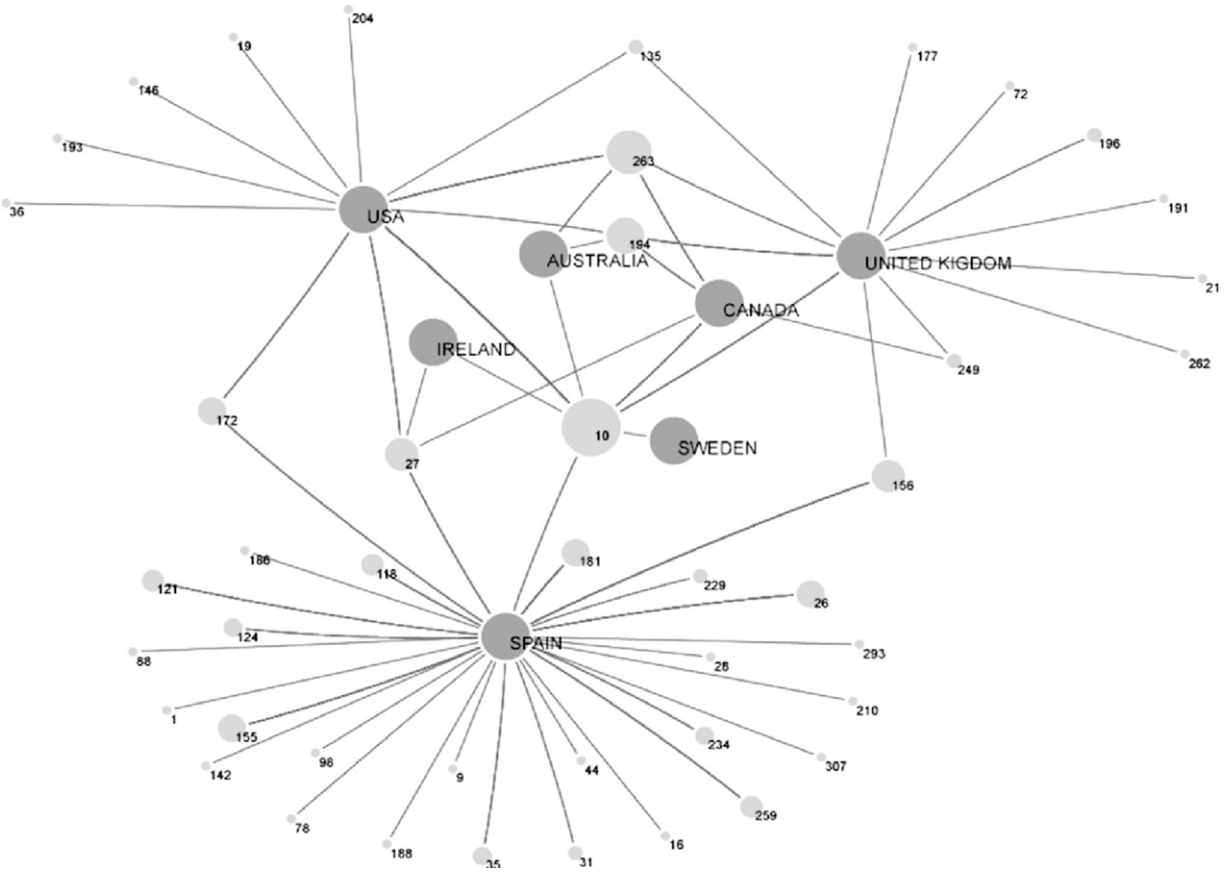

Fig. 9. Geographical despersion of book titles in Economics from the University of Navarra.

Table 15
Number of inclusions per library for book titles in Economics from the University of Navarra.

University-Library Catalog	Country	No. of titles	%Titles
Universidad Autonoma de Madrid	Spain	12	27
Universidad de Alcala Henares	Spain	11	25
Universidad Carlos III de Madrid	Spain	10	23
Consejo Superior de Investigaciones Cientificas	Spain	8	18
Universidad de Granada	Spain	8	18
Yale University	United States	7	16
Universidad Complutense de Madrid	Spain	6	14
Columbia University	United States	5	11
London School of Economics and Political Science	United Kingdom	5	11
King's College London	United Kingdom	4	9
Massachusetts Institute of Technology	United States	4	9
The University of British Columbia	Canada	4	9
Universidad de Burgos	Spain	4	9
University of Birmingham	United Kingdom	4	9

located in a particular country is represented in Fig. 9 by a link between the node corresponding to that title and the node representing that country. One can distinguish three groups. A first relates to titles that are only included in Spanish libraries (i.e., nodes 31, 44, 9, etc.). A second group consists of titles only included in US or UK libraries (i.e. nodes 204, 19, 72, 21, etc.) Finally, in the central area of the network there is a group of titles that are included in catalogs from various countries (i.e., nodes 263, 194, 10, etc.).

Finally, Table 15 shows the libraries that include titles published by UNAV researchers. UNAV book titles are particularly concentrated in four Spanish libraries located in Madrid. The library at the *Universidad Autónoma de Madrid* has the largest number of books and includes 27 percent of the total, while at the international level we find that *Yale University* includes 16 percent of the UNAV titles.

5. Discussion and conclusions

This paper explored the use of Library Catalog Analysis, defined as the application of bibliometric or informetric techniques to a set of library online catalogs, to describe quantitatively a scientific-scholarly field on the basis of published book titles. It focused on its value as a tool in studies of Social Sciences and Humanities, especially its cognitive structures, main book publishers and the performance of its actors. It proposed an analogy model between traditional citation analysis of journal articles and Library Catalog Analysis of book titles. It described a process of data collection from online catalogs and the creation and cleaning of an off-line database with book titles. It identified a number of technical problems and showed how these can be solved. And it showed a series of informative analyses of the database created in this way.

However, there are a number of limitations that must be taken into account in the interpretation of the results and, related to this, a number of issues and problems that need to be addressed and solved in a further development of the methodologies explored in this paper.

(i) National biases

Firstly, the set of libraries selected in the current study has a rather strong Anglo-Saxon *bias*. 70 percent of the libraries is located in the United States, United Kingdom, Canada and Australia. The set of libraries included in the current study cannot be assumed to constitute a sufficiently representative sample of major academic libraries in the world, especially in assessments of research performance in Social Sciences and Humanities.

This bias may especially affect the identification the national publishers active in countries that are not covered by the off-line database. The extent to which it affects the lists of the major international publishers in the field should be further examined. This research should also investigate in more detail the possible tendency that authors from a particular country publish books with publishers from the same country, and that academic libraries have a preference for purchasing books published from their own country.

The bias in geographic coverage is mainly caused by the fact that library catalog systems complying with the Z39.50 client–server protocol are not evenly distributed among countries. But collective catalogs like *WorldCat* or combined catalogs using *OCLC* could open new opportunities to include catalogs from countries not yet covered, and thus improve the degree of representativity of the set catalogs included in an analysis. A follow-up study should further explore these possibilities.

(ii) Field delimitation problems

A *second* limitation relates to the delimitation a scientific discipline by selecting specific words from subject headings. The current, exploratory study used the truncated string *Econom**. This string may not cover the entire discipline sufficiently well. For instance, book title records catalogued merely with the headings *Business* or *Finance* were not retrieved. This problem is particularly relevant in the assessment of research performance of institutions. This is one of the main reasons why only 19 percent of the book titles from the University of Navarra were found in the off-line database created

in the study. The search strings used to retrieve records should in some way be related to the thematic profile of the institution under assessment.

However, application of this principle may not be an easy task, since there may be differences in the subject classification systems among library catalogs. One and the same book title can be assigned to different subject headings in different catalogs, even within the same country. The implication is that if a title is not retrieved from a catalog using the search string *Econom** applied in this study, it does not follow that it is not included in that catalog; it may be included under a different subject heading. The extent to which the outcomes presented in this paper change in function of changes in the delimitation of the discipline awaits further research.

A related issue is the extent to which difference exist among the ways publishers classify their own books into broad subject groups, and the extent to which such classifications influence librarians in assigning their subject classifications to book titles. For instance, if may be a profitable marketing strategy for publishers to label their books in social science fields only weakly related to Economics as "Economics". If librarians are influenced by such a label, the outcomes presented in this paper may at least partly reflect publishers' marketing strategies.

A general solution is to delimitate in a first step a field in a library catalog in a rather broad way, so that the recall is sufficiently high. In a next step, selected records should be downloaded and an off-line database should be created. Finally, within this off-line database the field can be further delimited in a more accurate way, thus improving the precision. The use of keywords from the book titles should be further explored.

(iii) Types of libraries

It is important to take in account the orientation of a library. One can distinguish between two types of libraries: libraries with an educational orientation and those with a research orientation. These two types of libraries have different acquisition methodologies. In the first one the students and teachers play an important role. In the second one the librarian has a more important role and his or her decisions are crucial for the development of the collection. In LCA to be used in research assessment studies it may be appropriate to select catalogs from the research-oriented libraries only. The question as to how a library's orientation can be measured awaits further research.

(iv) Types of books

A *fourth* limitation relates to a classification of books into 'types'. Journal articles in large publication databases such as Thomson Scientifics' Web of Science are categorized into a number of document types, including normal article, review article and letter. But the academic library catalogs from which book titles were retrieved in the current study do not apply a categorization of books into types. Useful types would be: authored versus edited words; books primarily for teaching versus those primarily for research; books for a specialized scientific-scholarly audience versus those for broader audiences within or even outside the scientific-scholarly community. A follow-up study could deal with the development and implementation of such a useful, practical classification system.

(v) Analogy between LCA citation analysis

In the introduction section an analogy model was proposed between citation analysis of journal articles and Library Catalog Analysis of book titles. It was assumed that the number of catalogs in which a set of book titles is included gives an indication of the utility of those titles for the consumption of knowledge within the academic community, in the same way as citations to sets of documents are assumed to express the utility of those documents. This analogy has to be further explored and validated. One way of doing this is to examine in fields in which both journal and book publications are important, and in which the available citation databases (Thomson Scientific's Web of Science and Elsevier's Scopus) have a good journal coverage, – such as Economics –, the degree of correlation between the number of times book titles are cited in the serial literature on the one hand, and the number of library catalogs in which they are included on the other. This analysis could also take into account the geographic location of the citing authors and that of the libraries included in the LCA.

Despite these limitations and issues that need to be addressed in future studies, the study presented in this paper has shown that the proposed analogy model between citation analysis of journal articles and Library Catalog Analysis of book titles has proven to be valuable. Within the framework of this model a series of useful indicators and thematic mapping techniques was proposed and applied in the analysis of the database. It was illustrated how LCA can be fruitfully used to assess book production and research performance at the level of individual researchers, research departments, countries and book publishers. The outcomes show that Library Catalog Analysis of published book titles can be developed into a powerful and useful tool in studies of Social Sciences and Humanities.

Acknowledgements

The work described in this paper was carried out during a post doctoral stage of the first author at the Centre for Science and Technology Studies (CWTS). The authors are grateful to the University of Navarra and to Evaristo Jimenez-Contreras at the University of Granada for funding this stage. In addition, the authors wish to thank Dr. Janus Linmans (CWTS), Dr. Werner Mueller, Executive Vice President Business/Economics and Statistics at Springer, Heidelberg, and Dr. Rafael Olivares, librarian at the library of the University of Granada, for their useful comments on earlier drafts of this paper, and Prof. Rafa Bailon-Moreno at the University of Granada for creating the maps based on Copalred software.

References

Bailón-Moreno, R. (2003). Ingeniería del conocimiento vigilancia tecnológica aplicada a la investigación en el campo de los tensioactivos. Desarrollo de un modelo ciencimétrico unificado. Granada: Universidad de Granada. Departamento de Ingeniería Química.

Calhoun, K. (2007). The changing nature of the catalog and its integration with other discovery tools. http://www.loc.gov/catdir/calhoun-report-final.pdf.

Callon, M., Courtial, J. P., & Laville, F. (1991). Co-word analysis as a tool for describing the network of interactions between basic and technological research: The case of polymer chemistry. *Scientometrics*, *22*(1), 155–205.

Callon, M., Courtial, J. P., Turner, W. A., & Bauin, S. (1983). From translations to problematic networks: An introduction to co-word analysis. *Social Science Information*, *22*, 191–235.

Cherchye, L., & Vanden Abeele, P. (2005). On research efficiency. A micro analysis of Dutch university research in Economics and Business Management. *Research Policy*, *34*, 416–495.

Edward, S. (1999). Citation analysis as a collection development tool: A bibliometric study of polymer science theses and dissertations. *Serials Review*, *25*(1), 11–21.

Evans, P. (2001a). Z39.50: Part 1, overview. *Biblio Tech Review*. http://www.biblio-tech.com/html/z39_50.html.

Evans, P. (2001b). Z39.50: Part 2–technical details. *Biblio Tech Review*. http://www.biblio-tech.com/html/z39_50_part_2.html.

Glanzel, W., & Schoepflin, U. (1999). A bibliometric study of reference literature in the sciences and social sciences. *Information Processing and Management*, *35*, 31–44.

Hicks, D. (1999). The difficulty of achieving full coverage of international social science literature and the bibliometric consequences. *Scientometrics*, *44*(2), 193–295.

Lewison, G. (2001). Evaluation of books as research outputs in history medicine. *Research Evaluation*, *10*(2), 89–95.

Lindhom-Romantschuk, Y., & Warner, J. (1996). The role of monographs in scholarly communication: And empirical study of philosophy, sociology and economics. *Journal of Documentation*, *52*(4), 389–404.

Linmans, J. (2008). Een exploratieve studie van de onderzoeksprestaties van de Faculteit Letteren aan de Universiteit Leiden (in Dutch). Internal CWTS Report.

Moed, H. F. (2005). Implications for the use of the ISI citation indexes in research evaluation. In *Citation analysis in research evaluation*. Dordrecht (the Netherlands): Springer., p. 137–143.

Moed, H. F., De Bruin, R. E., & Van Leeuwen, T. (1995). New bibliometric tools for the assessment of national research performance. Database description, overview of indicators and first application. *Scientometrics*, *33*(3), 381–422.

Pancheshnikov, Y. (2007). A comparison of literature citations in faculty publications and student theses as indicators of collection use and a background for collection management at a University Library. *The Journal of Academic Librarianship*, *33*(6), 674–683.

Price, D. J. D. (1970). Citation measures of hard science, soft science, technology and non science. In C. E. Nelson & D. K. Pollock (Eds.), *Communication among scientist and engineers* (pp. 3–22). Lexington, MA: D.C. Heath.

Swygart-Hobaugh, A. J. (2004). A citation analysis of quantitative/qualitative methods debate's reflection in sociology research: Implications for library collection development. *Library Collections acquisitions & Technical Services*, *28*(2), 180–195.

Van Leeuwen, T. (2006). The application of bibliometric analyses in the evaluation of social science research. Who benefits from it, and why it is still feasible. *Scientometrics*, *66*(1), 133–154.

Villagrá Rubio, A. (1992). Scientific production of Spanish universities in the field of social sciences and language. *Scientometrics*, *24*(1), 3–19.

Wolfe Thompson, J. (2002). The death of the scholarly monograph in the Humanities? Citation patterns in literary scholarship. *Libri*, *52*(3), 121–136.