Information Processing and Management 45 (2009) 131–141

Contents lists available at ScienceDirect

Information Processing and Management

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

The integration of open access journals in the scholarly communication system: Three science fields

Tove Faber Frandsen

Royal School of Library and Information Science, Birketinget 6, 2300 Copenhagen S., Denmark

ARTICLE INFO

Article history:
Received 15 February 2008
Received in revised form 5 May 2008
Accepted 16 June 2008
Available online 8 August 2008

Keywords:
Open access journals
Bibliometrics
Field differences

ABSTRACT

The greatest number of open access journals (OAJs) is found in the sciences and their influence is growing. However, there are only a few studies on the acceptance and thereby integration of these OAJs in the scholarly communication system. Even fewer studies provide insight into the differences across disciplines. This study is an analysis of the citing behaviour in journals within three science fields: biology, mathematics, and pharmacy and pharmacology. It is a statistical analysis of OAJs as well as non-OAJs including both the citing and cited side of the journal to journal citations. The multivariate linear regression reveals many similarities in citing behaviour across fields and media. But it also points to great differences in the integration of OAJs. The integration of OAJs in the scholarly communication system varies considerably across fields. The implications for bibliometric research are discussed.

1. Introduction

Studies show that the influence of OAJs in the scholarly communication system is growing. Kling and Callahan (2003) provide an overview of studies on perception of OA journals. The study by McVeigh (2004) documents that the number of OAJs in the citation indexes provided by ISI Thomson™ is growing, both in terms of creating new titles and conversion of established titles. Furthermore, OAJs are dominantly lower-ranking journals in their field measured by Journal Impact Factor (JIF) and Immediacy Index although OAJs rank higher by Immediacy Index, than by JIF. Sotudeh and Horri (2007a) analyse the performance of OAJs in terms of expected citation rates and conclude that OA is widely recognised by scientific communities.

However, the influence of OAJs is not growing with the same rate in all fields as field differences are indicated in the existing literature in terms of the number and the acceptance of OAJs. The sciences are undeniably leaders in establishing OAJs, however, they are distributed unevenly within the sciences (Borgman, 2007, p. 186). In addition, Kling and McKim (2000) conclude that

> "[C]ommunicative plurality and communicative heterogeneity are durable features of the scholarly landscape, and [...] we are likely to see field differences in the use of and meaning ascribed to communications forums persist, even as overall use of electronic communications technologies both in science and in society as a whole increases." (Kling & McKim, 2000, p. 1306)

The study by McVeigh (2004) shows great field differences, and high-ranking OA journals are the most likely to be found in the fields of physics, engineering and mathematics.

The present study aims to contribute to the understanding of the implications of the increasing number of scientific publications published by journals running under an OA model. The influence of OAJs is typically measured by their number or

E-mail address: tff@db.dk

0306-4573/$ - see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.ipm.2008.06.001

share; however, there are no bibliometric studies on whether they are integrated or recognised generally in the scholarly communities. Another aspect yet to be investigated is whether the integration of OAJs in both OAJs and non-OAJs varies from field to field. More specifically, this paper addresses the following research question: are OAJs integrated in the journal communication system? The research question can be specified further: Does the citing of OA and non-OA journals depend on the citing journal being OA or non-OA and do field differences influence that citing behaviour?

The paper is structured as follows: The following section formulates an operational definition of OAJs to be used in this study. The next section presents the collected data and the chosen methods, followed by a presentation of the results of the analysis. Furthermore, the paper includes a discussion, and the last section contains conclusions and the perspectives of the paper.

2. Open access journals

Open access means that scientific publications are made freely available on the Internet, without any access restrictions. OA can be achieved using a number of different financing models. Thus, Willinsky (2006) identifies ten different models of providing open access to scholarly publications including both self-archiving and OAJs. OAJs can be seen as the second phase or strategy in the process of achieving open access (Brody & Harnad, 2005). The open access publishing strategy comprises of creating or converting traditional toll-access journals into open access journals. Furthermore, the strategy includes finding funding support for the publication costs and persuading authors to publish in OAJs. A few years ago, about 4% of scholarly journal titles and 1–2% of articles were directly published as open access (Harnad et al., 2004).

According to Moed (2007), the term open access is used with two different definitions. It is used to specify scientific publications published in a journal running under an open access model. However, it is also used to specify scientific publications that are freely available, not considering if they where originally published in a journal running under an OA model or in a journal managed under other business models but characterised by being deposited in a freely accessible archive such as a personal homepage, institutional repository or subject-based archive (Moed, 2007, p. 2047). This study focuses on OAJs; however, non-OAJs and OAJs can in practise be difficult to separate as non-OAJs can de facto be at least partially OA. Journals managed under other business models than OA can be partly OA because single publications can be self-archived by e.g. the author(s). Some journals provide free access after an embargo period. Furthermore, many scholars have full text access to non-OAJ articles through university or corporate licences implying that these scholars would not perceive non-OAJ and OAJ differently in terms of access.

The focus of this study is on a potential keenness or reluctance of authors (or editors) to integrate OAJs in the reference lists of the accepted publications in non-OAJs and OAJs. The key issue is not the accessibility of single publications but on the perception of quality tied to OAJs in various disciplines. Thus, self-archived non-OAJ articles do not distort the point made here. However, the vast field differences in use of OA and choice of OA model do necessitate that the fields are analysed separately.

The operational definition of OAJs in this study is journals managed under a business model that does not charge readers or their institutions for access. Journals with limited free access (e.g. free access is restricted to a select period of time or a select sample of publications in the journal) are not regarded as OAJs.[1]

3. Methods

A statistical analysis of the importance of type of media (OA versus non-OA) on the citing and cited side is an approach similar to the one used by Baldi (1997, 1998) on document level.[2] Using multiple linear regressions on both cited and citing journals enables controlling for different characteristics of the journals as well as for their degree of interaction or dependency. The method has been used by Frandsen (2005), albeit with a focus on the degree of dependency between American and European economics journals.

The journals included in this study were selected from three science disciplines. Ideally, social sciences and humanities disciplines could have been included but as relatively high numbers of OAJs within the selected disciplines are required, the present study is restricted to the sciences. The data in this study was analysed as three separate datasets as it is crucial to be able to control for discipline and sub-discipline specific variances. The analysis was performed on the basis of data from 2006 as this was the most recent publication year completed in the citation databases at the time of the data collection. The disciplines selected were biology, mathematics, and pharmacy and pharmacology as described by the classification scheme of Ulrich's Periodicals Directory™. They were selected on the basis of the number of OAJs which varies considerably among disciplines.[3] For the statistical analyses it is essential to select a discipline with a relatively high number of OAJs. A considerable number of OAJs convert into non-OAJs (Sotudeh & Horri, 2007b), however, as the development over time is not the focus in

[1] Ulrich's Periodicals Directory™ bases their distinction of OAJs and non-OAJs on a similar definition.

[2] Although it is remarkably complex to determine what citations measure (the reader is referred to Nicolaisen (2007) for recent review of theories of citation analysis), the citation analysis framework provides the opportunity to analyse what Cronin (2001, p. 2) refers to as "links (reference citations) provided routinely by authors in their reports and papers [which] are a means of exposing the underlying socio-cognitive structure of science."

[3] The number of OAJs is presently about 2700 according to the Directory of Open Access Journals (DOAJ, http://www.doaj.org). However, this number can be divided into disciplines showing significant differences in the number of OAJs.

this study it is sufficient that the journal had OA status at the time of data collection. An overview of the included journals and some of their characteristics is provided in Appendix 1–3. Within the discipline of biology journals from 5 biology sub-disciplines were selected resulting in 74 journals, within mathematics 25 general journals, and within pharmacy and pharmacology 20 journals from two sub-disciplines were selected. The smaller number of journals in the two latter disciplines is due to a smaller number of OAJs in these two disciplines and resulted in the exclusion of language as independent variable in those models.

The variables in the linear regressions consisted of a dependent variable and a number of independent variables. The dependent variable in this study was the dependency of the citing journal on the cited journal. Dependency was measured by the number of references from one journal to another which was determined using the Science Citation Index (SCI). The cited work field is uncontrolled and consequently attention must be paid to the different forms of names as well as articles in press. However, as the total number of references and the number of references to other publication types than journal articles (e.g. books and working papers) varies considerably across journals,[4] normalisation is needed. The number of references is normalised by dividing the number of references by the total number of references in the citing journal and multiplying by a hundred to get the relative dependency in per cent. The relative dependency of journal *i* on journal *j* is defined as follows:

$$\text{Relative dependency}_{i,j} = \frac{\text{Number of references}_{i,j} * 100}{\text{Number of references}_{i}}$$

The distribution of references across journals is expected to be influenced by a number of factors not related to the issue of OA. These factors were sought captured by a number of independent variables as they could potentially distort the results if not included. The relative dependency of journals was primarily described by the following independent variables: sub-discipline, JIF, publication patterns, OA and variables describing the relationship of the citing and cited journal. The sub-disciplinary variables consisted of dichotomous variables of the sub-disciplines (a journal could belong to more than one sub-discipline within the discipline as it depended on the indexing of the journals in Ulrich's). Geographical relations were described by a variable containing the geographic location of authors i.e. the share of authors located in three regions: North America, Western Europe and the rest of the world.[5] Furthermore, variables concerning the languages of the journals were constructed. A variable concerning the languages of the journals were constructed. Information on the language of an article was available in the citation indexes and a variable was created describing the language of the publication as being English or non-English. Information on the JIF and total number of citations received by the journals was available in Journal Citation Reports (JCR). It is essential to control for the average number of citations received per publication in any analysis of citing behaviour on journal level. Some journals receive 10 or 100 times as many citations as other journals and much higher degrees of dependencies on such journals must be expected. This is important as the focus of this study is not on explaining why some journals receive more citations than others, but rather to focus on the importance of the OA status of the cited journals in the citing journal. A variable describes the share of reviews (as journals consisting of many reviews are expected to be less likely to cite other journals also consisting of many reviews). The dichotomous OA variable was constructed on the basis of information from Ulrich's and confirmed on the journals' websites. Finally, a number of variables described the dyadic character of the relationship between journals. The variables sought to capture the effect of own group preference which is an effect detected on many levels. Self-citations is a strong own group preference as confirmed by e.g. Fassoulaki, Paraskeva, Papilas, and Karabinis (2000), Aksnes (2003), Frandsen (2005) and Frandsen (2007). Other, but probably weaker, own group variables are variables describing similarities between the citing and cited journal (e.g. same sub-discipline). The variable, dependency on this data sample, is the combined relative dependencies of a journal to all the journals in the dataset. The degree a journal depends on the other journals in this data sample should, in principle, increase the dependency on each single journal in the sample.

The results of the analysis presented below consist of different statistical analyses of the data material. One of the variables mentioned above was not included in the final models as it did not contribute to the understanding of the dependent variable (language). The slope coefficients for the linear relationships are given. Pearson's r^2 reveals information about the degree of correlation between the dependent and the independent variables when controlling for the effects of the other variables.

4. Results

In the following focus will be upon the variable characterizing whether the cited journal is an OAJ or not. If the coefficient to this variable in the statistical analysis is positive it indicates that this field has shown citing behaviour that gives more citations to OAJs than would have been expected on the basis of their characteristics. Thus such a field we shall describe as "*OA including*". If the coefficient to the variable is negative it oppositely indicates that journals in this field are citing OAJs less than their characteristics would imply, and such a field we shall denote as "*OA excluding*". Finally, if the coefficient is

[4] Some journals have more references to other document types such as monographs and working papers and an increased dependency on these document types should be reflected in lower dependencies on the journals in this data sample. The importance of the journal article is varying considerably across disciplines (Moed, 2005, pp. 129–130) and sub-disciplines (Frandsen & Nicolaisen, 2008).

[5] In this study, the general division of regions by the World Trade Organisation (WTO) is applied.

insignificantly different from zero, we shall denote the field as "*OA neutral*" as no apparent difference in the dependence of OA and non-OA can be found.

Table 1 shows the results of a multivariate linear regression analysis of the citing behaviour of all journals from the samples within the three disciplines. It is apparent that the citing behaviour of all three fields is relatively well described by such an analysis, as *R* squared exceeds one half in all three regressions.

From Table 1 is it clear that when looking at the disciplines as a whole we should denote biology and mathematics as OA neutral whereas pharmacy and pharmacology is OA excluding. Turning briefly to the other variables in the statistical models we see that there is a rather large degree of homogeneity in the importance of these across the three fields. In all three cases it is thus the case that the variables "Dependency on this data sample", "Cited journal JIF" and "Indicator for journal self-citations" influence the citing behaviour in the same (expected) way. It is thus to be expected that a higher JIF will tend making the degree of dependency higher. Similarly, the degree a journal depends on the other journals in this data sample should increase the dependency on each single journal in the sample – at least on average. Finally, it is well-known that a large percentage of citations are journal self-citations. In this context this translates into a higher degree of dependency on a journal when this is actually the journal itself. Since the dependency variable is measured in per cent, the coefficients to the indicator for self-citations show that the share of self-citing is between 1.2 and 1.7 percentage points higher than to a journal with otherwise similar characteristics. The variable "Combined share of reviews" is significant in two of the analyses with a negative sign. The reason for is that journals publishing a large share of review to a smaller extent are dependent on each other than on other types of journal where the reviewed literature is published. The importance of sub-discipline is also remarkable, although already established in the existing literature by e.g. Bordons and Zulueta (1997), Frandsen and Nicolaisen (2008).

Coming back to our main variable of interest, namely the OA indicator variable, the significant coefficient to this variable in the regression for pharmacy and pharmacology means that an average OAJ received eight citations less from each of the other journals than a similar non-OAJ in the year 2006. As already pointed out, this type of analysis cannot provide explanations of low or high levels of dependencies of a single journal across the data sample implying that the OAJs of OA including and excluding fields probably receive relatively more citations from journals not included in the study or outside their field. The focus of the study is on the analysis of OAJs and non-OAJs separately.

In order to analyze whether the citing behaviour is different for OAJs and non-OAJs, an extra two sets of multivariate regressions have been carried out, where the OAJs and non-OAJs have been analyzed separately. The results of these analyses are shown in Tables 2 and 3.

From Table 2, it is apparent that when we look only at traditional non-OAJs we still find that pharmacy and pharmacology is OA excluding and biology is OA neutral. However, contrary to the overall status of mathematics, non-OAJs in this field are OA including, giving more citations to OAJs in the sample than to other similar journals. In terms of numbers this means that the non-OAJs in mathematics gave 2 citations extra to each of the OAJs in the sample. The coefficients to the control variables are almost identical to those in Table 1, confirming the overall robustness of the method.

From Table 3 we can see that for OAJs the analysis gives rise to somewhat different results than the two previous tables. We thus see that within biology OAJs are OA including whereas OAJs within mathematics, and pharmacy and pharmacology apparently are neutral with respect to their citing behaviour towards other OAJs. The coefficient to the OA variable for biology corresponds to one extra citation to each OAJ from each OAJ compared to the number of citations to a similar non-OA journal.

Table 4 summarizes the findings with respect to citation behaviour towards OAJs for the three analyzed disciplines. As was already apparent in the presentation of the statistical analyses above there are great field differences in the integration of OAJs

Table 1
Multivariate linear regression analysis of citing behaviour of all journals

Variable	Biology	Mathematics	Pharmacy and pharmacology
Intercept	−0.0140	−0.3175	
Cited journal OA			−0.1078
Dependency on this datasample	0.0085	0.0492	0.0506
Cited journal JIF	0.0094	0.1825	0.0336
Share of authors from Western countries	−0.0307		
Indicator for journal self-citations	1.2037	1.7419	1.4534
Difference in JIF between citing and cited journal	−0.0013		
Belong to same sub-discipline	0.0528	–	0.1381
Combined share of reviews	−0.0407		−0.1884
R squared	0.504	0.536	0.542
# of observations	5476	625	400

Numbers not shown are not significant at the .05 significance level. The sign "–" indicates that this variable was not included in the analysis for Mathematics.
The dependent variable is relative dependency in per cent.

Table 2
Multivariate linear regression analysis of citing behaviour of non-OAJs

Variable	Biology	Mathematics	Pharmacy and pharmacology
Intercept	−0.0187	−0.3453	
Cited journal OA		0.1193	−0.1165
Dependency on this datasample	0.0100	0.0496	0.0468
Cited journal JIF	0.0077	0.2002	0.0316
Share of authors from Western countries	–0.0329		
Indicator for journal self-citations	1.3578	2.1000	1.2963
Difference in JIF between citing and cited journal			
Belong to same sub-discipline	0.0498	–	0.1218
Combined share of reviews	−0.0423		−0.1630
R squared	0.560	0.596	0.598
# of observations	3404	400	240

Numbers not shown are not significant at the .05 significance level. The sign "–" indicates that this variable was not included in the analysis for Mathematics.
The dependent variable is relative dependency in per cent.

Table 3
Multivariate linear regression analysis of citing behaviour including OAJs

Variable	Biology	Mathematics	Pharmacy and pharmacology
Intercept		−0.3159	
Cited journal OA	0.0248		
Dependency on this datasample	0.0050	0.0485	0.0481
Cited journal JIF	0.0231	0.1545	0.0840
Share of authors from Western countries			
Indicator for journal self-citations	0.9237	1.1151	1.6710
Difference in JIF between citing and cited journal	−0.0148		
Belong to same sub-discipline	0.0612	–	0.1684
Combined share of reviews	−0.0517		−0.2948
R squared	0.428	0.473	0.517
# of observations	2072	225	160

Numbers not shown are not significant at the .05 significance level. The sign "–" indicates that this variable was not included in the analysis for Mathematics.
The dependent variable is relative dependency in per cent.

Table 4
Summary of OA inclusion in three sciences and subdivisions hereof

	Overall	Non-OA	OA
Biology	0	0	+
Mathematics	(+)	+	0
Pharmacy and pharmacology	−	−	0

+ indicates OA inclusion, − indicates exclusion while 0 indicates neutrality. Sign shown in parentheses indicates significance at the .1 significance level.

The main conclusion to be drawn from the table thus seems to be that any statement indicating a uniform advantage or disadvantage for OAJs is questionable. In fact the analysis in this paper seems to imply that the status of OAJs and the subsequent citations to them in OAJs as well as non-OAJs depend greatly on the fields and subfields in question.

5. Discussion

Before addressing the consequences of the findings it must be emphasized that the analysis in this paper only has included a subset of the science disciplines. The results of the empirical study are based on references from three science disciplines and cannot necessarily be generalized to other fields. Furthermore, self-archiving makes the demarcation of OAJs and non-OAJs vague. Finally, it should be noted that this type of analysis cannot provide explanations of low or high levels

of dependencies of a single journal across the whole data sample as this effect (to a large extent) is captured by the variable containing JIF values. The method can, however, explain lower or higher dependencies of a set of journals within a field. However, although limited, the results have implications for bibliometric studies.

To a wide extent OAJs and non-OAJs can be described by the same elements. Many of the same variables in the multiple linear regressions are statistically significant with identical signs providing evidence of the strength of the models used in this study. Both OAJs and non-OAJs cite journals with a high JIF more and journals depending greatly on this data sample have higher dependencies themselves as cited journals. Regardless of being OA or NOA the journals have strong own group dependencies in terms of self-citations and sub-discipline self-citations. These variables are included as control variables and expected to turn out statistically significant with a positive coefficient. However, the results of the analyses of OAJs and non-OAJs differ in terms of the use of OAJs depending on the discipline.

The present study contributes to the understanding of the so-called open access postulate defined as "authors are more likely to read, and thus cite, articles that are made available under an OA model" (Craig, Plume, McVeigh, Pringle, & Amin, 2007). The findings in this study indicate that if such an effect exists for OAJs it is probably not found in all disciplines. The development of OA is not just a matter of the number of OAJs in a field but also to what extent they are accepted and used in non-OAJs as well as OAJs.

This study gives insight into the developments in scholarly communication. As pointed out by Gläser (2003) the important issue is to what extent new forms of social order emerges due to the Internet. The use of Internet can be positively related to author productivity (Kaminer & Braunstein, 1998; Barjak, 2006), the Internet has facilitated large-scale collaborations (Finholt, 2002) and new communication regimes in biology based on online databases (Hilgartner, 1995). However, Gläser (2003) argues that the Internet rapidly creates new social phenomena but they are not necessarily sociologically new. The social structure of the scientific communities could remain unchanged although it reforms the work practices. As Van Raan (1997, p. 447) states:

> "In our opinion, the new electronic publishing developments will not influence conceptually [the] main functions of scientific communication. Of course, technology will certainly influence, even dramatically, these functions, particularly in terms of performance and of mechanical improvement."

Furthermore, Van Raan (2001, p. 63) argues that "Plus ça change, plus c'est la même chose". There are examples of the Internet not necessarily changing social phenomena. Lorigo and Pellacini (2007) have shown steady and constant growth in the frequency of long distance scholarly collaborations in a physics community and Mackenzie Owen (2007) finds that OAJs does not transform the research article by incorporating specific digital properties. It is complex to identify the new forms of social order emerging due to the Internet and thus separating them from new social phenomena that are not sociologically new. Following Barjak (2006) there are two major concerns:

- Causation.
- Distinguishing between function and technology.

It must be stressed that one cannot make any causal arguments on the basis of the present analysis as it can point to the underlying structure of OAJs and non-OAJs in various fields but not explain why there are varying patterns of interaction among journals. It could be an issue of author perceptions of OA. Findings by Swan and Brown (2004) show that the main reason for not having published in an OAJ is lacking familiarity of the concept of OAJs or with specific OAJs in their field. Authors who have not published in OAJs perceive them to have low prestige and impact, directly in contrast to the perception of authors who have published in an OAJ. Another possible explanation is the specific research areas within sub-disciplines as Zhao (2005) and Talja et al. (2005) find publishing behaviour being closely related to the research area of the author. A third perspective is the issue of access. Authors publishing in OAJs could have less access to articles published in non-OAJs and they must thus depend more on the publications freely available on the Internet in e.g. OAJs. However, it could also be self-archived publications which are beyond the scope of this analysis to investigate.

On the other hand the second concern is easily determined as this study is based on data material with the same function. Both OAJs and non-OAJs are publishing mediums and the difference in citing behaviour is thus to be found in the financing models or the culture surrounding the journals. Summing up, it should be emphasized that until the causality has been further investigated one should be careful making definite conclusions on the integration of OAJs in the scholarly communication system. However, it can be concluded that there are differences in the scholarly communication in OAJs and non-OAJs across fields.

The results have implications for all researchers conducting bibliometric studies and the consequences for bibliometric research could be widespread. It will affect individuals or groups under study, how the data source used for the evaluation is covering OAJs and non-OAJs. Bibliometric studies can be made using a wide variety of data sources and perhaps a combination of several data sources is preferable (Zhao & Strotmann, 2007). Regardless, the choice of data source(s) the relative share of OAJs can make a difference for the evaluation of individuals or institutions as it is related to a different citing behaviour within some fields. Should a pool of documents be sampled for further analysis, it is obviously of great importance how this is done. Such a sample could be problematic in terms of the relative size of OAJs and non-OAJs represented in the sample and one should be careful making generalizations (Nicolaisen, 2006).

To ensure valid results of bibliometric studies it is crucial to recognize possible biases in coverage in terms of the access aspect that may lead to biased results. In an evaluation that takes place across a wide board of journals (being both OA and non-OA) attention must be paid to the factors that may be determining the results of the analysis and appropriate precautions must be taken before initiating bibliometric studies using journal articles from either one or both as pools of documents. Analyses into the underlying structures of a discipline provide valuable insight in the scholarly communication of that field. Journal interaction analysis as performed in this study can reveal some of the hidden structures that are determinants for the results from citation analysis.

6. Conclusion

The results of this study confirm the many similarities between OAJs and non-OAJs which are in accordance with the resemblance in function. However, the results also point to dissimilarities. Causation cannot be determined in the present study; however, it is clear that in some fields authors publishing in OAJs are demonstrating different citing behaviour than authors publishing in non-OAJs. Within biology the non-OAJs are OA neutral with respect to their citing behaviour towards other OAJs. The OAJs within biology are OA including, giving more citations to OAJs in the sample than to other similar journals. Within pharmacy and pharmacology the non-OAJs as well as the OAJs are OA excluding, giving less citations to OAJs in the sample than to other similar journals. Finally, within mathematics non-OAJs are OA including whereas OAJs are neutral. Even within OAJs there is no guarantee of acceptance and integration of OAJs in general on the level we would expect based on a comparison with non-OAJs with similar characteristics.

Acknowledgement

The author is greatly indebted to Jeppe Nicolaisen and Birger Hjørland. Furthermore, the author acknowledges the valuable comments by an anonymous referee.

Appendix 1. Biology journals included in the study

	OA	Dependency on this data sample	Share of reviews	Share of authors from North America	Share of authors from Western countries
Acta Biochimica Polonica	1	4.0	0.1	0.1	0.2
Acta Bioquimica Clinica Latinoamericana	1	1.9	0.1	0.0	0.1
Acta Protozoologica	1	3.2	0.0	0.0	0.5
Acta Zoologica	1	0.8	0.1	0.0	0.3
Advances in Biochemical Engineering–Biotechnology	0	1.9	0.0	0.9	1.0
Advances in Carbohydrate Chemistry and Biochemistry	0	1.2	0.7	0.1	0.7
African Zoology	0	1.3	0.0	0.1	0.4
American Journal of Hematology	0	2.6	0.0	0.4	0.6
American Journal of Primatology	0	2.7	0.0	0.7	0.8
American Museum Novitates	1	4.0	0.0	0.8	0.8
Animal Genetics	0	2.8	0.0	0.2	0.7
Annual Review of Genetics	0	2.1	1.0	0.5	1.0
Applied Biochemistry and Microbiology	0	4.2	0.0	0.0	0.0
Archives of Biochemistry and Biophysics	0	4.4	0.0	0.5	0.8
Archives of Microbiology	0	6.2	0.0	0.2	0.6
Behavior Genetics	0	2.3	0.0	0.5	0.9
Biochemistry	0	3.8	0.0	0.6	0.8
Biological Chemistry	0	4.0	0.1	0.3	0.8
Bioprocess and Biosystems Engineering	0	5.1	0.0	0.2	0.5
Bioscience, Biotechnology, and Biochemistry	1	5.5	0.0	0.0	0.0
BMC Biotechnology	1	6.2	0.0	0.2	0.8
BMC Evolutionary Biology	1	3.6	0.0	0.3	0.8

(*continued on next page*)

Appendix 1 (*continued*)

	OA	Dependency on this data sample	Share of reviews	Share of authors from North America	Share of authors from Western countries
BMC Genetics	1	3.4	0.0	0.4	0.8
BMC Genomics	1	5.0	0.0	0.3	0.8
BMC Microbiology	1	5.5	0.0	0.3	0.7
Brazilian Archives of Biology and Technology	1	10.1	0.0	0.0	0.1
Brazilian Journal of Microbiology	1	4.8	0.0	0.0	0.1
Bulletin of the American Museum of Natural History	1	1.2	0.9	0.4	0.6
Canadian Journal of Microbiology	0	4.9	0.0	0.4	0.6
Caribbean Journal of Science	1	1.5	0.0	0.6	0.6
Clinical Biochemistry	0	2.0	0.1	0.4	0.6
Clinical Microbiology and Infection	0	3.5	0.1	0.1	0.8
Contributions to Zoology	1	2.2	0.1	0.1	0.8
Current Microbiology	0	7.4	0.0	0.1	0.3
Electronic Journal of Biotechnology	1	5.6	0.1	0.0	0.1
Food Microbiology	0	4.0	0.0	0.3	0.7
Genes & Genetic Systems	1	4.8	0.0	0.0	0.1
Genes, Brain and Behavior	0	1.9	0.2	0.4	0.8
Genetics and Molecular Biology	1	4.3	0.0	0.1	0.1
Genome	0	4.4	0.0	0.3	0.6
IEEE Transactions on Information Technology in Biomedicine	0	1.4	0.0	0.3	0.7
Indian Journal of Biochemistry and Biophysics	0	4.4	0.1	0.0	0.0
International Journal of Molecular Sciences	0	1.8	0.0	0.1	0.2
International Microbiology	1	5.6	0.4	0.3	0.9
Journal of Animal Ecology	0	1.6	0.0	0.3	0.8
Journal of Basic Microbiology	0	5.6	0.0	0.1	0.3
Journal of Biochemistry and Molecular Biology	1	4.2	0.1	0.1	0.1
Journal of Biomedicine and Biotechnology	1	2.9	0.6	0.5	0.9
Journal of Chemical Technology and Biotechnology	0	4.0	0.0	0.1	0.5
Journal of Clinical Microbiology	0	5.4	0.0	0.3	0.7
Journal of Genetics	1	3.7	0.0	0.0	0.3
Journal of Lipid Research	1	3.4	0.1	0.5	0.8
Journal of Microbiology and Biotechnology	0	7.5	0.0	0.1	0.1
Journal of Molecular Catalysis B: Enzymatic	0	6.6	0.0	0.1	0.3
Journal of Plant Biochemistry and Biotechnology	0	5.5	0.0	0.1	0.1
Journal of Proteome Research	0	4.8	0.0	0.4	0.8
Korean Journal of Genetics	0	6.8	0.0	0.0	0.1
Laboratory Animals	0	4.4	0.1	0.2	0.8
Microbiology and Immunology	1	4.1	0.0	0.1	0.1
Microbiology and Molecular Biology Reviews	0	2.2	1.0	0.4	0.8
Molecular and Cellular Biology	0	3.4	0.0	0.1	0.4
Molecular Biology	0	5.1	0.1	0.0	0.1
Mutagenesis	0	3.3	0.1	0.1	0.8
Nature Biotechnology	0	6.3	0.0	0.6	0.9
North American Journal of Aquaculture	0	2.1	0.0	0.9	1.0
Nucleic Acids Research	1	5.3	0.0	0.4	0.8

Appendix 1 (*continued*)

	OA	Dependency on this data sample	Share of reviews	Share of authors from North America	Share of authors from Western countries
Pathobiology	0	1.2	0.1	0.1	0.5
Process Biochemistry	0	6.2	0.0	0.1	0.3
Raffles Bulletin of Zoology	1	3.3	0.0	0.2	0.3
Russian Journal of Genetics	0	4.2	0.1	0.0	0.1
Steroids	0	3.6	0.0	0.2	0.6
Trends in Biotechnology	0	4.3	0.7	0.3	0.8
Trends in Microbiology	0	4.1	0.9	0.4	0.9
Zoosystema	1	2.0	0.1	0.1	0.5

Appendix 2. Mathematics journals included in the study

	OA	Dependency on this data sample	Share of reviews	Share of authors from North America	Share of authors from Western countries
Applied Mathematics and Computation	0	3.6	0.0	0.1	0.1
Annales Academiae Scientiarum Fennicae	1	7.3	0.0	0.2	0.6
Annals of Mathematics	1	7.3	0.0	0.5	0.4
Bulletin of the American Mathematical Society	1	3.9	0.1	0.6	0.3
Communications in Algebra	0	8.0	0.0	0.2	0.3
Discrete Mathematics	0	6.5	0.0	0.3	0.3
Duke Mathematical Journal	0	7.3	0.0	0.4	0.5
Electronic Communications in Probability	1	1.2	0.0	0.3	0.5
Electronic Research Announcements in Mathematical Sciences	1	2.7	0.0	0.4	0.3
Electronic Transactions on Numerical Analysis	1	2.1	0.0	0.4	0.4
European Journal of Applied Mathematics	0	2.6	0.0	0.1	0.5
Forum Mathematicum	0	6.7	0.0	0.3	0.5
Houston Journal of Mathematics	0	7.3	0.0	0.4	0.2
Izvestiya Mathematics	0	4.5	0.0	0.0	0.1
Journal of Nonlinear Mathematical Physics	1	2.9	0.0	0.1	0.3
Journal of the American Mathematical Society	0	7.1	0.0	0.6	0.3
Journal of the London Mathematical Society	0	7.2	0.0	0.2	0.5
Mathematical Logic Quarterly	0	3.6	0.0	0.2	0.4
Mathematical Problems in Engineering	1	1.6	0.0	0.1	0.1
Mathematical Research Letters	0	7.2	0.0	0.5	0.3
Mathematical Social Sciences	0	3.1	0.0	0.2	0.5
NODEA - Nonlinear Differential Equations and Applications	0	2.9	0.0	0.0	0.7
Proceedings of the Japan Academy. Series A.	1	5.1	0.1	0.0	0.1
Quarterly Journal of Mathematics	0	8.1	0.0	0.2	0.5
Studies in Applied Mathematics	0	3.4	0.0	0.4	0.3

Appendix 3. Pharmacy and pharmaceutical journals included in the study

	OA	Dependency on this data sample	Share of reviews	Share of authors from North America	Share of authors from Western countries
AAPS Journal	1	3.3	0.6	0.9	0.1
AAPS PharmSciTech	1	8.6	0.0	0.3	0.2
American Journal of Pharmaceutical Education	1	4.5	0.0	0.9	0.0
Biological & Pharmaceutical Bulletin	1	3.8	0.0	0.0	0.0
British Journal of Clinical Pharmacology	0	4.5	0.0	0.1	0.3
Chemical & Pharmaceutical Bulletin	1	5.3	0.0	0.0	0.1
Clinical and Experimental Pharmacology & Physiology	0	1.9	0.1	0.2	0.1
Current Pharmaceutical Biotechnology	0	0.5	0.9	0.3	0.4
Formulary	0	1.2	0.0	0.9	0.0
Journal of Pharmaceutical and Biomedical Analysis	0	4.8	0.0	0.1	0.3
Journal of Pharmacology and Experimental Therapeutics	0	4.1	0.0	0.5	0.3
Journal of Pharmaceutical Sciences	0	8.0	0.0	0.5	0.3
Journal of Pharmacy and Pharmaceutical Sciences	1	5.0	0.2	0.4	0.1
Journal of Physiology and Pharmacology	1	4.4	0.0	0.1	0.2
Molecular Pharmacology	0	3.7	0.0	0.5	0.3
Pharmaceutical Biology	0	3.1	0.0	0.1	0.1
Pharmaceutical Research	0	5.8	0.1	0.4	0.3
Pharmacological Reports	1	3.4	0.1	0.1	0.1
Pharmacology	0	3.5	0.0	0.1	0.5
Trends in Pharmacological Sciences	0	2.6	0.7	0.3	0.5

References

Aksnes, D. W. (2003). A macro study of self-citations. *Scientometrics, 56*(2), 235–246.

Baldi, S. (1997). *A network approach to the analysis of citation flows: A comparative study of two research areas in the natural and the social sciences*. Colombus: Department of Sociology, Ohio State University, Colombus.

Baldi, S. (1998). Normative versus social constructivist processes in the allocation of citations: A network analytic model. *American Sociological Review, 63*, 829–846.

Barjak, F. (2006). The role of the Internet in informal scholarly communication. *Journal of the American Society for Information Science and Technology, 57*(10), 1350–1367.

Bordons, M., & Zulueta, M. A. (1997). Comparison of research team activity in two biomedical fields. *Scientometrics, 40*(3), 423–436.

Borgman, C. L. (2007). *Scholarship in the digital age: Information, infrastructure, and the Internet*. Cambridge, MA: MIT Press.

Brody, T., & Harnad, S. (2005). Keynote lecture: Providing open access to peer-reviewed articles to maximise and measure their research impact. <http://www.oai.unizh.ch/symposium/docs/Harnad.ppt>.

Craig, I., Plume, A., McVeigh, M., Pringle, J., & Amin, M. (2007). Do open access articles have greater citation impact? A critical review of the literature. *Journal of Informetrics, 1*(3), 239–248.

Cronin, B. (2001). Bibliometrics and beyond: some thoughts on web-based citation analysis. *Journal of Information Science, 27*(1), 1–7.

Fassoulaki, A., Paraskeva, A., Papilas, K., & Karabinis, G. (2000). Self-citations in six anaesthesia journals and their significance in determining the impact factor. *British Journal of Anaesthesia, 84*(2), 266–269.

Finholt, T. (2002). Collaboratories. *Annual Review of Information Science and Technology, 36*, 73–107.

Frandsen, T. F. (2005). Journal interaction: a bibliometric analysis of economics journals. *Journal of Documentation, 61*(3), 385–401.

Frandsen, T. F. (2007). Journal self-citations – Analysing the JIF mechanism. *Journal of Informetrics, 1*(1), 47–58.

Frandsen, T. F., Nicolaisen, J. (2008). Intra-disciplinary differences in database coverage and the consequences for bibliometric research. *Journal of the American Society for Information Science and Technology*.

Gläser, J. (2003). What Internet use does and does not change in scientific communities. *Science Studies, 16*(1), 38–51.

Harnad, S., Brody, T., Vallieres, F., Carr, L., Hitchcock, S., Gingras, Y., et al. (2004). The green and the gold roads to open access. Nature (web focus). <http://www.nature.com/nature/focus/accessdebate/21.html>.

Hilgartner, S. (1995). Biomolecular databases: New communication regimes for biology? *Science Communication, 17*(2), 240–263.
Kaminer, N., & Braunstein, Y. M. (1998). Bibliometric analysis of the impact of internet use on scholarly productivity. *Journal of the American Society for Information Science, 49*(8), 720–730.
Kling, R., & Callahan, E. (2003). Electronic journals, the internet, and scholarly communication. *Annual Review of Information Science and Technology (ARIST), 37*, 127–177.
Kling, R., & McKim, G. (2000). Not just a matter of time: Field differences in the shaping of electronic media in supporting scientific communication. *Journal of the American Society for Information Science, 51*(14), 1306–1320.
Lorigo, L., & Pellacini, F. (2007). Frequency and structure of long distance scholarly collaborations in a physics community. *Journal of the American Society for Information Science and Technology, 58*(10), 1497–1502.
Mackenzie Owen, J. (2007). *The scientific article in the age of digitization*. Dordrecht: Springer.
McVeigh, M. E. (2004). Open access journals and the ISI citation database: Analysis of impact factors and citation patterns. Thomson Scientific Whitepaper. <www.thomsonisi.com/media/presentrep/essayspdf/openaccesscitations2.pdf>.
Moed, H. F. (2005). *Citation analysis in research evaluation*. Dordrecht, NL: Springer.
Moed, H. F. (2007). The effect of 'Open Access' upon citation impact: An analysis of ArXiv's condensed matter section. *Journal of the American Society for Information Science and Technology, 58*(13), 2047–2054.
Nicolaisen, J. (2006). Traditional author co-citation analysis: A discussion of the sampling problem. In *Proceedings of the international conference on multidisciplinary information sciences and technologies, InSciT2006* (pp. 635–639).
Nicolaisen, J. (2007). Citation analysis. *Annual Review of Information Science and Technology, 41*, 609–641.
Sotudeh, H., & Horri, A. (2007a). The citation performance of open access journals: A disciplinary investigation of citation distribution models. *Journal of the American Society for Information Science and Technology, 58*(13), 2145–2156.
Sotudeh, H., & Horri, A. (2007b). Tracking open access journals evolution: Some considerations in open access data collection validation. *Journal of the American Society for Information Science and Technology, 58*(11), 1578–1585.
Swan, A., & Brown, S. N. (2004). JISC/OSI journal authors survey report. <http://www.jisc.ac.uk/uploaded_documents/JISCOAreport1.pdf>. [downloaded 30.7.07].
Talja, S., Savolainen, R., & Maula, H. (2005). Field differences in the use and perceived usefulness of scholarly mailing lists. *Information Research, 10*(1) [paper 200].
Van Raan, A. F. J. (1997). The future of the quality assurance system: its impact on the social and professional recognition in the era of electronic publishing. *Journal of Information Science, 23*(6), 445–450.
Van Raan, A. F. J. (2001). Bibliometrics and internet: Some observations and expectations. *Scientometrics, 50*(1), 59–63.
Willinsky, J. (2006). *The access principle: The case for open access to research and scholarship*. Cambridge, MA: MIT Press. XV, 287p.
Zhao, D. (2005). Challenges of scholarly publications on the web to the evaluation of science: A comparison of author visibility on the web and in print journals. *Information Processing and Management, 41*(6), 1403–1418.
Zhao, D., & Strotmann, A. (2007). Can citation analysis of web publications better detect research fronts? *Journal of the American Society for Information Science and Technology, 58*(9), 1285–1302.