Journal of Informetrics 9 (2015) 79–89

Contents lists available at ScienceDirect

Journal of Informetrics

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

Features of top-rated gold open access journals: An analysis of the scopus database

Gianfranco Ennas*, Maria Chiara Di Guardo

Department of Economics & Business, University of Cagliari, Viale S. Ignazio 17, 09123 Cagliari, Italy

ARTICLE INFO

Article history:
Received 5 July 2014
Received in revised form 3 October 2014
Accepted 17 November 2014
Available online 15 December 2014

Keywords:
Gold open access ranking
Open access impact analysis
SJR
Open access journals

ABSTRACT

Purpose: The goal is to identify the features of top-rated gold open access (OA) journals by testing seven main variables: languages, countries, years of activity and years in the DOAJ repository, publication fee, the field of study, whether the journal has been launched as OA or converted, and the type of publisher.
Sample: A sample of 1910 gold OA journals has been obtained by combining Scopus SJR 2012, the DOAJ, and data provided by previous studies (Solomon, 2013).
Method: We have divided the SJR index into quartiles for all journals' subject areas. First, we show descriptive statistics by combining quartiles based on their features. Then, after having converted the quartiles into a dummy variable, we test it as a dependent variable in a binary logistic regression.
Contribute: This work contributes empirically to better understanding the gold OA efficacy of data analysis, which may be helpful in improving journals' rankings in the areas where this is still a struggle.
Findings: Significant results have been found for all variables, except for the types of publishers, and for born or converted journals.

1. Introduction

Research quality has always played a crucial role for scholars, publishers, professional societies, and funding organizations (Falagas, Kouranos, Arencibia-Jorge, & Karageorgopoulos, 2008). Authors compete for the opportunity to publish their research in high-quality and highly ranked journals in order to gain the largest diffusion possible. Research impact is the degree to which findings are read, used, applied, built-upon, and cited by users in their own further research and applications; again, it is a measure of the progress and productivity of studies (Harnad et al., 2004). Understanding journals' performances in terms of the impacts they have is a significant challenge; bibliometrics works in this field, developing and studying indexes and indicators and providing statistics of various types. In the last decade, open access (OA) has become an established and well-known phenomenon, and the number of journals and articles released in OA has grown rapidly. Although OA has not changed how research is conducted (Pinfield, 2005), it has upset the rules of publishing scholarly articles. A great amount of literature has attempted to measure OA's success and efficacy by looking at its metrics since the origin of OA (McVeigh, 2004).

In particular, previous works have studied some features (see Tables 1 and 2 for a detailed list) of gold OA (where the publisher provides free online access) journals' rankings, which are intended as an impact factor or other similar indicator, in

* Corresponding author. Tel.: +39 3460246997.
E-mail address: gianfrancoennas@unica.it (G. Ennas).

http://dx.doi.org/10.1016/j.joi.2014.11.007
1751-1577/

Table 1
Highlights of previous studies.

Author(s)	Purpose	Data source	Method	Variables	Principal findings
McVeigh (2004)	Overview of OA journals in Thomson Reuters database.	ISI Web of Knowledge DOAJ J-STAGE ScIELO	Descriptive statistics and graphics	Journals by subject Regions Impact Factor, Immediacy Index, Percentile Rank Journals' age	High-ranking OA journals are not evenly distributed among the subject fields. Geographical distribution differs strongly from the ISI citation databases as a whole. Two-thirds of the journals are below the 50th percentile in rank. The average number of years in OA varies a bit between fields. OA publishing is growing slowly, both by creation of new titles and by conversion.
Giglia (2010)	Test the performance of OA journals with the most traditional bibliometric indicators.	JCR DOAJ	Descriptive statistics and graphics	Geographical distribution Macro disciplinary areas, Impact Factor, percentile, Immediacy index, 5-year Impact Factor Journals' age	Low presence of OA journals in JCR 2008. Nearly 71% of OA in JCR 2008 Science edition comes from Central and South America. IF performance is 38.62% in the 0–50 percentiles. The direct causal relationship between age and visibility and prestige in terms of citations cannot be straightforwardly inferred.
Miguel, Chinchilla-Rodriguez, and de Moya-Anegón (2011)	H1: There is a stronger gold road in the social sciences and humanities. H2: There is a greater proportion of gold road journals in emerging economic regions. H3: These journals are not the most relevant and are, for the most part, from the last quartile.	SCOPUS DOAJ Sherpa-RoMEO	Descriptive statistics, graphics and Chi-square test	Gold, green and no OA Main Field of Study SJR Index Geographic region	The gold road has a greater proportion of work in journals belonging to Medicine; Biochemistry, Genetics and Molecular Biology; Areas Related to Medicine; and Earth and Environmental Sciences. The peripheral and emerging regions have greater proportions of gold road texts, and publications in green journals are almost nonexistent. The gold road journals have no visibility and, for the most part, belong to the fourth quartile regardless of the geographic area of origin.
Björk and Solomon (2012)	Comparing the scientific impact of OA journals with subscription journals, controlling for journal age, the country of the publisher, discipline and their business model.	Combining Datasets: Ulrichsweb, SJR, DOAJ	Descriptive statistics and graphics	SJR, Countries, Journal launch period, Publication Fee	Seventy percent of subscription journals are from publishers in the four major publishing countries. Newer journals, particularly in medicine and health, are performing at about the same level as subscription journals. The funding mechanism of a journal is irrelevant in considering its quality. Almost half of OA journals started before 1996 and were not published in top publisher countries.
Gumpenberger, Ovalle-Perandones, and Gorraiz (2013)	To identify the number of Gold Open Access journals that have successfully taken the hurdle to be indexed in JCR and allocated to disciplines, countries and quartiles. To analyze the temporal evolution of the gold OA journals' impact.	Ulrichsweb DOAJ JCR SJR CWTS	Descriptive statistics, graphics, and Linear regression Correlation	IF, SJR, SNIP, Countries, Timeline, Publisher distribution	The impact of top Gold OA journals is generally increasing for all of the analyzed indicators. One-third of the newly launched OA titles were already indexed in JCR after 1 year, and 80% of these received an IF at least within a 5-year interval. The percentage of Q1 titles is <20% in the UK, and the USA contributes to 80% of the top Gold OA. Gold OA is still small compared to the total number of scholarly journals worldwide. Q1 titles are predominantly journals in English.

Table 1 (*Continued*)

Author(s)	Purpose	Data source	Method	Variables	Principal findings
Solomon, Laakso, and Björk (2013)	(1) Document the growth of OA journals and articles between 1999 and 2010. (2) Compare SNIP2 Citation Averages for OA journals with that of subscription journals during the period of 1999 through 2010.	SJR and DOAJ	Descriptive statistics and graphics	SNIP2, Journal Born or converted as OA, Publication fee, Sciences (health non health) Language of converted journals	High quality OA publishing is growing at a rapid rate and in the case of OA publishing, funded by APCs. There is evidence based on citation rates that these journals are on par with subscription journals. The bulk of converted OA journals are located outside the four major publishing countries. A high percentage of the journals that are converted to OA are also in languages other than English.
Solomon (2013)	Types of organizations and their characteristics of OA journals' publishers in Scopus.	SJR and DOAJ	Descriptive statistics and graphics	Publisher Type, Discipline, Publication Fee	Professional, society, and university publishers own approximately 85% of the journals and articles published. Over 80% of the journals published by professional publishers charge APCs. There appears to be no clear delineation between funding models and types of publishers. The types of organizations publishing OA journals differed significantly across disciplines.

Table 2
Features studied by previous works.

	Ranking	Subject	Geographical distribution	Journals' age	Born/Converted as OA	Publication fee	Type of publisher	Language	Total
McVeigh	x	x	x	x					4
Giglia	x	x	x	x					4
Miguel et al.	x	x	x						3
Gumpenberger	x	x	x	x				x	5
Björk and Solomon	x		x	x		x			4
Solomon et al.	x	x			x	x		x	5
Solomon		x				x	x		4
Total	6	6	5	4	1	3	1	2	

comparison to other forms of OA and to traditional publishing methods. Of course, these studies have helped to understand the OA phenomenon, but they present two limitations: First, single variables are combined with the journals' rankings and show, for most cases, only the descriptive statistics. Second, each work usually focuses on a small set of variables to achieve its purpose; therefore, understanding the features of top journals as a whole is not easy because they use different data-sets built in different years. The aim of this paper is to understand what the relevant features of top gold OA journals are, and, in doing so, we use several descriptive statistics and a binary logistic regression model. With this type of regression method, on the one hand, we are able to classify the journal features by their relationship with the ranking; yet, on the other hand, we can identify the features that a journal should have to raise its ranking. The main motivation is related to the fact that a high ranking can be interpreted as a success factor for a journal; therefore, understanding what is a determinant for reaching success can contribute to the improvement of OA efficacy in the field of studies where OA is still struggling.

We have built a sample of 1910 gold OA journals, and we have used the journals' ranking as a dependent variable. In doing so, we have divided gold OA journals by the SJR index for each subject category. We have considered top journals (Q1) as those where the relative location is in the top 25% of the SJR distribution. We have decided to combine three databases; the main regard is Scopus because it contains a larger number of gold OA journals with an impact factor larger than those of other journals. It also includes titles from more countries that are published in a greater variety of languages (Leydesdorff, de Moya-Anegón, & Guerrero-Bote, 2010). Again, we have used the Directory of Open Access Journals (DOAJ), repository data, and other variables' information from previous studies' datasets (Solomon, 2013).

This paper is divided into six sections: In Section 2, we provide a literature review. In Section 3, we present the material and methods used. In Section 4, we present variables with some descriptive statistics and comparisons to other works. In Section 5, we discuss results of regression. Finally, in Section 6, we present our conclusions.

2. Background

A scientific publication represents the final stage of many months and sometimes years of meticulous planning, execution, and analyses of hundreds of experiments (Benos et al., 2005). With the advent of the Internet, more and more researchers are

making their research openly accessible by self-archiving it online to increase their visibility, usage, and citation impact. The publication of scientific content has been one of the areas to benefit most from the emergence of the Internet (Björk, 2004). OA had its first formal definition and guidelines in 2002, with the Budapest Open Access Initiative, followed by the Bethesda statement on open access publishing in 2003 and the Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities (2003), which is usually known as the BBB definition. OA can be defined as free and unrestricted access on the public Internet to literature that scholars provide without expectation of direct payment (Prosser, 2003). According to these definitions, publishing in OA means satisfying two conditions: The first is granting unrestricted access to anyone via the internet and the license to copy, use, distribute for non-commercial purposes, and make and distribute derivative works without any payments or restrictions. The aim is to remove barriers to literature in order to accelerate research, enrich education, and share learning. The second condition is to deposit a complete version of the work and all supplemental materials immediately upon initial publication in at least one online repository, including a copy of the permission in a suitable standard electronic format. Repositories consist of a physical space reserved for permanent or intermediate storage of archival material and can be searched and retrieved for later use (Hayes, 2005). Actually, the DOAJ, which is managed by Lund University, is the largest repository, including scientific journals presenting quality controls, that allows free access. OA occurs in two variants (Harnad et al., 2004): Green OA refers to publishing in any appropriate traditional journal, in addition to self-archiving the pre- or post-print paper in a repository. Gold OA refers to articles in fully accessible OA journals. The gold model uses a traditional journal publication system, and nothing is paid by the reader of a peer-reviewed article. Some journals require a fee, paid by the author's organization or the research funder and sometimes by the author. According to the DOAJ repository, in November 2013, there were 6573 journals that did not require a processing charge and 2652 that required a processing charge.

Regarding the benefits of the OA articles in terms of citations, there is not agreement in the literature. Early studies have claimed that OA articles are cited more often and published in less time in comparison to traditional publications. In particular, Lawrence (2001) reports that OA articles in computer science are cited more, and multidisciplinary works between diverse fields of study at varying stages of adoption of OA have confirmed that OA articles have a greater research impact than articles that are not freely available (Antelman, 2004). In a longitudinal study of a cohort of OA and non-OA articles with an article-level approach, it has been demonstrated that there is direct and strong evidence for preferential or earlier citation of articles published originally as OA (Eysenbach, 2006). However, some authors are critical about the cause and effect relationship between OA and higher citations, stating that the benefits of self-archiving may be uncertain and may vary between different fields of study (Craig, Plume, McVeigh, Pringle, & Amin, 2007). Again, Moed (2007) has highlighted two points: the first is a *self-selection* bias, that is, authors tend to self-archive high-quality articles and thus receive more citations. The second is that many works do not take into consideration a wide time windows to evaluate the benefits of OA in comparison to traditional publishing methods.

As can be noted in Table 1, during the last decade, several scholars have measured the impact of gold OA journals and have analyzed the main characteristics of them in comparison to green OA journals and to traditional journals. The common strategy has been to combine two or more databases. The first published study is a decade old (McVeigh, 2004) and uses the Thomson Reuters database. The results highlight that top OA journals are not equally distributed between different fields of study, with a prevalence of Physics, Engineering and Mathematics. McVeigh (2004) has also noted that over 55% of journals allow self-archiving; in regard to the geographical distribution, over one-third of OA journals were published in Asia-Pacific, while North America and Western Europe account for approximately 40% of OA titles. Again, she has shown that the overall mean percentile rank in terms of Journal Impact Factor was 39.8 percentile, while two-thirds of the journals were below the 50th percentile in rank. The mean percentile rank of OA journals by Immediacy Index was the 46 percentile. As a group, journals that have adopted an OA distribution model have not achieved a significantly greater citation impact. However, individual OA journals have been appearing among the highest ranked journals, even within a few years of their launch. More recently, Giglia (2010) has found a low presence of OA journals in JCR 2008, and she has confirmed, with some different results, that there are strong differences between disciplinary areas and impact, considering the best performances: in Medicine, there is a strong presence in the top twenty percentiles, 15.96%; 14.42% in Life Sciences; 12.63% in Mathematics, Physics, and Engineering; and 4.66% in Chemistry. Again, she has shown that there are many titles that rank low by Impact Factor but high by Immediacy Index. Regarding geographical aspects, she has also confirmed that nearly 71% of OA in JCR 2008 Science edition have come from Central and South America. Finally, she has stated a direct causal relationship between age, visibility and prestige, in terms of citations that cannot be straightforwardly inferred. Subsequently, Miguel, Chinchilla-Rodriguez, and de Moya-Anegón (2011: p. 1142) confirmed previous studies, but, regarding the impact, they found that, "*For the most part, belong to the fourth quartile regardless of the geographic area of origin.*" In contrast, Gumpenberger, Ovalle-Perandones, and Gorraiz (2013) have shown that gold OA journals' IF is increasing, and one-third of newly launched journals are indexed in JCR after a year. Again, in that work, they have shown a percentage lower than 20% of the journals in the first quartile and a concentration of 80% of the top gold OA journals in the UK and in the USA. Björk and Solomon (2012) have introduced another element to investigate: a comparison between OA journals that require publication fees (APCs) and journals that do not require payment. The result was that a funding mechanism is not related to the journals' quality. In the same paper, they have also shown that 70% of subscription journals are owned by organizations located in the four major publishing companies (the USA, the UK, the Netherlands and Germany). These results have been confirmed in the work of Solomon, Laakso, and Björk (2013), in which the Source-Normalized Impact per Paper Version 2 (SNIP2) for OA journals was combined with and without APCs during the period from 1999 through 2010.

They have found that, regardless of the business model, articles are cited at a similar rate to subscription journals. Again, converted OA journals have a SNIP2 lower than born OA journals, although some differences exist between health science and other fields of study. Again, they have found that a high percentage of converted journals are not published in the English language.

Finally, Solomon (2013) has introduced a classification between types of publishers (APCs) and platforms. Some results state that over one-third of the journals and 42% of the articles are owned by professional publishers, and the APCs are closely related to them. Moreover, in the same work, he has confirmed that a great number of OA journals across all disciplines are published outside the four major publishing countries.

In Table 2, we have summarized, in the first column, author(s) and, in the other columns, variables studied: ranking field of study, geographical distribution, journals' age and publication fee are the most recurrent. In the following paragraphs, we present our sample and the variables and compare our data to previous studies.

3. Material and method

The dataset consists of 1910 gold OA journals. In building it, the first step was to decide how to compare journals' rankings. In fact, the problem concerning indicators to evaluate the research is longstanding; several metrics exist to measure the ranking of a journal, and the most known is the Thomson Reuters IF. It has been criticized for multiple reasons (Bornmann, Marx, Gasparyan, & Kitas, 2012; Seglen, 1997); the most critical reasons for our study are as follows: (i) small research fields tend to lack journals with high impact; (ii) citation rates of articles determine journals' impacts but not vice versa; (iii) IF is a function of the number of references per article in the research field; and (iv) journal impact factors are not statistically representative of individual journal articles. In other words, different scientific areas, fields and micro-fields of study have different citation habits (Lancho-Barrantes, Guerrero-Bote, & Moya-Anegón, 2010). Recently, the Scopus SJR indicator began an alternative to Thomson Reuters IF; it also takes into account the "quality" of citations received by a journal, whereas the journal IF considers incoming citations only in a quantitative manner. Thus, the use of the SJR indicator allows for the estimation of a journal's impact, reducing the influence of self-citations such that they cannot be more than 33% of the total because prestige can be transferred to a journal by all other journals, but not by itself (González-Pereira, Guerrero-Bote & Moya-Anegón, 2010). Nonetheless, the introduction of the SJR indicator does not bring about radical changes in this regard. It has been shown that half of the journals in the IF top 100 journal list are placed within a reasonable range of 32 ranking places in the SJR indicator journal list (Falagas, Kouranos, Arencibia-Jorge, & Karageorgopoulos, 2008). A common workaround to overcome index limitations is to consider each journal in its field, determine the quartile or the percentile, and then compare them with all of the journals. Using this strategy, it is possible to benefit from the advantage of working with categorical variables. Other scholars have preferred to use a normalized impact factor, the SNIP index; however, in the literature, there are many criticisms about that indicator (Leydesdorff & Opthof, 2010; Leydesdorff, Radicchi, Bornmann, Castellano, & Nooy, 2013; Lee & Shin, 2014).

Having said this, we have used three sources. The first is the Scopus database; the total number of journals in Scopus is 20,544. SJR (SCImago 2007) *divides sciences into 310 subject areas and 27 subject fields. We have taken the database of a*ll journals by subject area, and then we have merged the 307 data-sheets (three were empty). In doing so, we have obtained the countries' variables and the quartile division. It is frequent that a journal is indexed in multiple subject areas and categories. To have manageable data, we have maintained individual journals in the highest quartile per subject area; for example, if the journal is in the second quartile in Nursing and in the third quartile in Health Professions, we have kept only Nursing; if it is in the same quartile, we have kept both. After that, we have merged the datasheet with the DOAJ dataset to identify gold OA and to obtain the following variables: publication fee, languages, age of the journals and the period they were added into the DOAJ repository. The third source is Solomon's database (Solomon, 2013), which is used to obtain the publisher type classification and to determine if the single journal has been born as OA or if it has been converted. Finally, we have cleaned the data, deleting a total of 100 observations because 45 journals have ended their activity before the end of 2012 and 55 were without quartile information. It is also important to note that this study covers all active gold OA journals in Scopus and over 20.70% of OA journals in the DOAJ.

4. Variables

4.1. Dependent variable

4.1.1. Ranking

The total number of journals in the first quartile is 288, the number in the second quartile is 503, the number in the third quartile is 652, and the number in the fourth quartile 467. We have converted the journals' quartile variable into a dichotomic variable called ranking, assigning 0 for quartiles 2, 3 and 4 and 1 for quartile 1. Hence, the value 0 is for 1622, and the value 1 is for 288 journals. Only approximately 15% of the journals are in the first quartile. Regarding the first quartile, similar results have been provided by Gumpenberger et al. (2013), in contrast to what was stated by Miguel et al. (2011).

Table 3
Top publishing languages.

	English	Multilingual	Spanish	Portuguese	French	German
1st Quartile	284	24	10	13	8	6
2nd Quartile	488	108	49	53	16	6
3rd Quartile	579	192	132	94	30	16
4th Quartile	367	177	149	62	25	21
Total	1718	501	340	222	79	49

Table 4
Top publishing languages in non-multilingual journals.

	English only	Spanish only	Portuguese only	French only	German only
1st Quartile	262	0	2	0	0
2nd Quartile	381	8	5	0	0
3rd Quartile	392	30	19	0	3
4th Quartile	207	52	7	1	8
Total	1242	90	33	1	11

4.2. Independent variables

Below, all of the independent variables used for our model are presented; descriptive statistics are shown, where possible, with all quartile information in order to give additional advice.

4.2.1. Language

The publishing language has been taken into consideration by Solomon et al. (2013) in reference to converted journals and by Gumpenberger et al. (2013) in reference to Q1 journals. In agreement with Lobachev (2008), we think that understanding the diversity of the information universe represents an important point for determining current trends in global information production. Hence, we want to investigate these variables to find additional insight. Almost all of the titles in our sample, 1718 out of 1910 and 284 out of 288 regarding the top-ranked journals, are in English or in English plus some other languages, and 501 journals publish in more than one language. The top languages in overall worldwide scholarly production are English, German, Chinese, Spanish and French (Lobachev, 2008). It is interesting to see that Chinese does not appear in the top positions. Conversely, Portuguese is not only at the third position but also has 13 journals in the first quartile (Table 3).

If we consider non-multilingual journals, English not only dominates, but, as seen in Table 4, other languages have a marginal role; just two journals in Portuguese are in the first quartile.

4.2.2. Country

Over 49% of the total number of journals and over 74% of the top ranked journals are concentrated in six countries. While the UK and the USA own 38.19% and 22.92% of the total number of top ranked journals (288), respectively, others hold only a small percentage. Geographical distribution has been studied by McVeigh (2004), Giglia (2010), and Miguel et al. (2011), but only statistics related to the continents have been shown. Björk and Solomon (2012) have analyzed the four major traditional publishers (the US, the UK, the Netherlands and Germany) versus others. The only paper showing statistics per country was produced by Gumpenberger et al. (2013); they compared the number of journals in Q1 in Ulrichsweb and in the DOAJ. With close results in both databases, they found the US to be in first place, the UK in second, Germany in third, Switzerland and Japan in fourth, and Canada in fifth. According to the DOAJ, the top publishing OA countries, in terms of the number of journals, are the United States with 12.37%, Brazil with 9.35%, the UK with 6.29%, India with 6.08%, and Spain with 5.31%. In the table below are the top publishing countries according to our sample (Table 5).

In Table 6, the top 10 publishing countries according to the Scopus database are compared, and the percentage denotes the total number of journals present on 12/31/2012 (the entire database contains 20,554 entries). In comparison to those results, we can see the top 10 gold OA publishing countries according to our sample and the DOAJ repositories. Both in our sample and in the DOAJ repository, the Netherlands and Germany, which are known as large publishers, are not in the first 10 positions. If we compare results from our sample to the DOAJ, we can say that the top five publishing countries are very close.

Table 5
Top publishing countries.

	USA	UK	Brazil	India	Spain	Japan	New Zealand
1st Quartile	66	110	12	6	5	7	7
2nd Quartile	71	80	58	35	22	23	21
3rd Quartile	76	20	79	44	37	22	25
4th Quartile	50	9	35	10	24	16	13
Total	263	219	184	95	88	68	66

Table 6
Comparison of top publishing countries in Scopus and DOAJ.

Scopus	Journals	%	Sample	Journals	%	DOAJ	Journals	%
USA	5605	27.28	USA	263	13.77	USA	1201	12.37
UK	5036	24.51	UK	219	11.47	Brazil	908	9.35
The Netherlands	1706	8.30	Brazil	184	9.63	UK	611	6.29
Germany	1213	5.90	India	95	4.97	India	590	6.08
China	538	2.62	Spain	88	4.61	Spain	516	5.31
France	487	2.37	Japan	67	3.51	Egypt	440	4.53
Japan	459	2.23	New Zealand	66	3.46	Germany	333	3.43
Italy	401	1.95	Turkey	62	3.25	Romania	297	3.06
Spain	393	1.91	Chile	60	3.14	Italy	287	2.96
India	369	1.80	Poland	59	3.09	Canada	262	2.70

Table 7
Type of publishers.

	Profess.	Society	Univ.	Indep.	Gov.	Other	Unknown	Total
1st Quartile	157	54	33	8	13	20	3	288
2nd Quartile	202	140	77	6	32	35	11	503
3rd Quartile	162	187	189	16	29	48	21	652
4th Quartile	107	128	160	6	17	26	23	467
Total	628	509	459	36	91	129	58	1910

4.2.3. Type of publishers

We have used Solomon's classification, and almost 84% of the journals are owned by three categories of publishers: Professional, Society and University. Hereinafter, we present the relative percentage of journals in the first quartile for all categories of publishers: 25.00% professionally published journals, 10.61% society journals, 7.19% university published journals, 22.22% independent scholar publishers, 14.29% government agencies, 15.50% other organizations, and 5.17% unknown publishers. However, independent scholar publishers possess only 36 journals, and the percentage in the first quartile is very close to that of professionally published journals (Table 7).

Regarding the temporal variables, we have tested the age of the single journal intended as the period between its foundation and 12/31/2012, as well as the years in the DOAJ repositories. McVeigh (2004) has found that journals adopting an OA distribution model have not achieved a significantly greater citation impact. More recently, Giglia (2010: p. 33) has stated: "*Direct causal relationship between age and visibility and prestige in terms of citations cannot be straightforwardly inferred*." Finally, Solomon et al. (2013) have found that the distribution ages for subscription and free journals are roughly equal. Giving diverse results by previous studies seems to be an important aspect to investigate for this variable, as is understanding if the permanence in the DOAJ repositories can influence the ranking.

Years of activity is a continuous variable referring to how many years the journal has existed and been published. The min is 1, the max is 132, the mean is 12.09, and σ is 10.24.

Years in DOAJ is a continuous variable and reports how many years the journal has been indexed in the DOAJ repository. The min is approximately six months, the max is 10 years and 7 months, the mean is 6.19, and σ is 2.65.

Born or converted is a Boolean variable that is assigned a value of 0 if the journal was born as OA (864) and a value of 1 (1030) if it has been converted. Although most of the journals have been converted, over 68% of the journals in the first quartile were born as OA journals. For 16 journals, information is missing. This variable has been studied only by Solomon et al. (2013), as they correctly noted that OA experiences growth in two distinct ways: first, by conversion of existing journals and, second, by the birth of new journals. As widely demonstrated, OA articles are, in general, cited more; however, as we can see in the table below, it seems that converted journals tend to obtain a low ranking (Table 8).

4.2.4. Publication fee

Distribution of this variable has a strategic importance because "pay to publish" might suggest a deterrent for gold OA. In this regard, Solomon et al. (2013) have demonstrated that journals without APCs have increased more rapidly than others. Unexpectedly, the number of articles published with APCs is higher than articles without a publication fee. This means that journals adopting a business model that requires a form of payment tend to publish a larger number of papers. Hence,

Table 8
Born or converted journals.

	Born	Converted	Total
1st Quartile	197	91	288
2nd Quartile	238	261	499
3rd Quartile	247	398	645
4th Quartile	182	280	462
Total	864	1030	1894

Table 9
Publication fee.

	No Pub fee	Pub fee	Total
1st Quartile	96	177	273
2nd Quartile	260	219	479
3rd Quartile	445	172	617
4th Quartile	327	120	447
Total	1128	688	1816

investigation of the APCs journals and their relationship with the ranking is interesting. According to our data, despite the fact that 62.11% of the journals do not require a publication fee, 61.45% of the top ranked journals do require one (Table 9).

4.2.5. Subject area

As previously stated, journals are frequently classified in more than one subject area; in our database, 475 out of 1910 journals are in more than one. Hence, the total number of journals per subject area is 2518 instead of 1910. There is an enormous difference between the percentage of top OA journals in the first quartile and the subject area, varying from 0% in Dentistry to over 44% in Energy. Again, in the last column, we can see the percentage of gold OA journals in the entire Scopus database; here, the percentage varies from 2.45% in Energy to 28.40% in Multidisciplinary (Table 10).

5. Results and discussion

Results of the logistic regression are shown in Table 11; the model has fit the data quite well. In fact, the chi-square test rejects the hypothesis of no explanatory power, and the model correctly predicted 83% of the observations. As expected, the subject area of Dentistry was automatically dropped by the software (STATA 13) because estimation is not possible when a covariate does not vary within the category of an independent variable (Freese & Long, 2006). In fact, ln (0) is undefined; that is, the variable's distribution does not permit a finite coefficient, and, therefore, this does not bias the remaining coefficients in the model. The software automatically checks for multi-correlation with no evidence of it. Further tests have been performed, and they reported a mean variance inflation factor (VIF) of 1.88; again, the condition number reported, 26.58, is below the critical value of 30, so we can say that the regression has non-significant multi co-linearity. To evaluate the effects of independent variables, the significance values (p) were analyzed. If the significant values are less than 0.05 (95% confidence interval), it can be said that the independent variables have an effect on the ranking. With caution, we can say that a positive regression coefficient means that the explanatory variable increases the probability of the outcome, while a negative regression coefficient means that the variable decreases the probability of that outcome (Table 12).

Table 10
Fields of study.

Subject area	OA	Q1	% on Q1	Total in Scopus	% of OA
Agricultural and Biological Sciences	247	22	8.91%	1692	14.60%
Arts and Humanities	74	12	16.22%	2102	3.52%
Biochemistry, Genetics and Molecular Biology	201	32	15.92%	1553	12.94%
Business, Management and Accounting	26	1	3.85%	854	3.04%
Chemical Engineering	35	1	2.86%	439	7.97%
Chemistry	42	3	7.14%	563	7.46%
Computer Science	84	8	9.52%	1084	7.75%
Decision Sciences	8	1	12.50%	178	4.49%
Dentistry	14	0	0.00%	119	11.76%
Earth and Planetary Sciences	81	16	19.75%	853	9.50%
Economics, Econometrics and Finance	36	1	2.78%	640	5.63%
Energy	9	4	44.44%	367	2.45%
Engineering	109	7	6.42%	1893	5.76%
Environmental Science	82	15	18.29%	889	9.22%
Health Professions	23	4	17.39%	218	10.55%
Immunology and Microbiology	61	7	11.48%	458	13.32%
Material Science	52	6	11.54%	706	7.37%
Mathematics	72	10	13.89%	963	7.48%
Medicine	705	133	18.87%	5478	12.87%
Multidisciplinary	23	6	26.09%	81	28.40%
Neuroscience	48	10	20.83%	337	14.24%
Nursing	32	7	21.88%	371	8.63%
Pharmacology, Toxicology and Pharmaceutics	91	12	13.19%	554	16.43%
Physics and Astronomy	51	5	9.80%	568	8.98%
Psychology	43	3	6.98%	717	6.00%
Social Sciences	236	27	11.44%	3413	6.91%
Veterinary	33	3	9.09%	177	18.64%
Total	2518	356		27,267	

Table 11
Logistic regression results.

	Odds ratio	Std. error	*Z*	*P*> |z|	[95% Conf. Interval]	
English*	6.1936530	3.7279520	3.03	0.002	1.9037490	20.150420
Multilingual*	0.2527771	0.1095607	−3.17	0.002	0.1080944	0.591116
Spanish**	0.3705111	0.1776948	−2.07	0.038	0.1447339	0.948489
Portuguese	2.3790430	1.3972540	1.48	0.140	0.7524483	7.521906
French	1.2528960	0.6785869	0.42	0.677	0.4334026	3.621915
German	2.9442690	1.7766910	1.79	0.074	0.9022522	9.607867
USA*	1.9782060	0.4611308	2.93	0.003	1.2527140	3.123859
UK*	5.3925550	1.3765240	6.6	0.000	3.2697430	8.893559
Brazil	1.2027000	0.6430883	0.35	0.730	0.4217124	3.430031
India	0.5466591	0.2600571	−1.27	0.204	0.2151703	1.388835
Spain	1.6865990	1.0016790	0.88	0.379	0.5265983	5.401871
Japan	0.7066355	0.3845526	−0.64	0.523	0.2432034	2.053152
New Zealand	0.7741365	0.3522113	−0.56	0.574	0.3173538	1.888389
Turkey	0.2353511	0.2462435	−1.38	0.167	0.0302775	1.829417
Poland	0.6641993	0.4281880	−0.63	0.526	0.1877397	2.349853
Years of activity	1.0115100	0.0082510	1.4	0.161	0.9954667	1.027811
Years in DOAJ*	1.0904830	0.0361491	2.61	0.009	1.0218850	1.163687
Pub Fee*	2.0841300	0.4892931	3.13	0.002	1.3154900	3.301887
Agricultural and Biological Sciences	1.4097910	0.4248569	1.14	0.254	0.7809715	2.544921
Arts and Humanities*	6.0241560	2.6731210	4.05	0.000	2.5245810	14.374840
Biochemistry, Genetics and Molecular Biology	1.2153530	0.3374231	0.7	0.482	0.7053104	2.094232
Business, Management and Accounting	0.3556165	0.3914741	−0.94	0.348	0.0411104	3.076183
Chemical Engineering	0.4376612	0.4640878	−0.78	0.436	0.0547696	3.497327
Chemistry	1.1955240	0.8105124	0.26	0.792	0.3165783	4.514767
Computer Science	0.9307175	0.4061676	−0.16	0.869	0.3956865	2.189195
Decision Sciences	0.9017095	1.0897060	−0.09	0.932	0.0844110	9.632391
Earth and Planetary Sciences*	3.6979950	1.4398090	3.36	0.001	1.7240600	7.931957
Economics, Econometrics and Finance	0.3561789	0.3806080	−0.97	0.334	0.0438617	2.892349
Energy*	23.3672500	20.3649200	3.62	0.000	4.2342110	128.956400
Engineering	0.6220425	0.2978551	−0.99	0.321	0.2433525	1.590026
Environmental Science	1.3922260	0.5216103	0.88	0.377	0.6680350	2.901485
Health Professions	3.4356430	2.3889130	1.77	0.076	0.8793066	13.423810
Immunology and Microbiology	0.6372469	0.3066982	−0.94	0.349	0.2481053	1.636739
Material Science	1.6366410	0.9283956	0.87	0.385	0.5383982	4.975116
Mathematics	1.1482720	0.5166916	0.31	0.759	0.4753651	2.773720
Medicine**	1.7590100	0.3951165	2.51	0.012	1.1325770	2.731926
Multidisciplinary*	4.4293140	2.5456210	2.59	0.010	1.4359280	13.662810
Neuroscience**	2.6651550	1.2406300	2.11	0.035	1.0702560	6.636774
Nursing	1.9013670	1.0975180	1.11	0.266	0.6133817	5.893879
Pharmacology Toxicology and Pharmaceutics	1.5658660	0.6049308	1.16	0.246	0.7343711	3.338826
Physics and Astronomy	0.4956044	0.3217324	−1.08	0.280	0.1388542	1.768932
Psychology	1.7591720	1.2358480	0.8	0.421	0.4439409	6.970945
Social Science*	2.3886400	0.7482052	2.78	0.005	1.2927800	4.413434
Veterinary	1.2798170	0.8721241	0.36	0.717	0.3365899	4.866250
Born or Converted	0.7322029	0.1509078	−1.51	0.130	0.4888751	1.096642
Professionally published Journal	0.9218220	0.6249184	−0.12	0.904	0.2441200	3.480893
Society Journals	1.3829560	0.9093563	0.49	0.622	0.3811575	5.017785
University published journals	1.0703100	0.7224771	0.1	0.920	0.2850589	4.018692
Independent scholar publisher	1.5080790	1.2398400	0.5	0.617	0.3010398	7.554823
Government Agency	2.6355750	1.9610160	1.3	0.193	0.6131033	11.329670
Other Organization	1.8591310	1.3077210	0.88	0.378	0.4683509	7.379870
Constant	0.0048070	0.0043770	−5.86	0.000	0.0008069	0.028000

* Statistically significant at 1% significance level.
** Statistically significant at 5% significance level.

Table 12
Measures of fit for logistic regression.

Log-Lik Intercept Only:	−764.173	Log-Lik Full Model:	−581.789
D (1747)	1163.578	LR(51):	364.768
		Prob>LR:	0.000
McFadden's R^2	0.239	McFadden's Adj R^2:	0.171
ML (Cox–Snell) R^2:	0.184	Cragg-Uhler (Nagelkerke) R^2:	0.321
McKelvey & Zavoina's R^2:	0.425	Efron's R^2:	0.233
Variance of y:	5.722	Variance of error:	3.29
Count R^2:	0.863	Adj Count R^2:	0.092
AIC:	0.705	AIC* n:	1267.578
BIC:	−11,930.163	BIC':	17.476
BIC used by Stata:	1553.317	AIC used by Stata:	1267.578

Regarding the first variable, Language, we can say that English is highly and positively related to the ranking; however, this is not surprising news, as we are studying scientific journals, and English is recognized as the "standard language" to share research. In Section 3, we have shown that over a quarter of journals are published in more than one language; according to our results, it is interesting to see how this strategy is strongly and negatively related to the ranking. If we look more deeply at the data, we can observe that 476 out 501 multilingual journals contain English as a publishing language, hence it is possible to state with evidence that English is a discriminant, but only if the journal is not multilingual. Again, we have strong negative presumption about Spanish; for other languages, there are no signs of significance. The second category of analyzed variables is the country, as we have already seen in the descriptive statistics; the larger OA publisher countries are quite different in comparison to the entire Scopus database. Looking at results, we can say that journals owned by UK and US publishers have a very strong and positive relation to the ranking; other countries do not show signs of significance. Gumpenberger et al. (2013) show a higher rate of success for Japanese JCR indexed OA titles; the number of journals in their study appears coherent with our sample, but the results are not significant. A temporal variable yields an important outcome because an answer in this regard has not yet clearly been given by previous studies. With the results of regression, it seems clear how journals benefit from the increase of the permanence in the DOAJ repositories, while the age of the journal has no significant relation with the ranking. Regarding publication fee, from previous studies (Solomon, 2013), we already know that APCs journals are lower in comparison to others. We can say with a very strong presumption (p value lower than 1%) that journals adopting a business model requiring a form of payment to publish tend to become top rated more than others. This result is in contrast to the work of Björk and Solomon (2012), which states that a funding mechanism is not related to the journal's quality. The fifth category regards the field of study, as seen by looking at regression results; the situation is very varied, and an outcome of 7 out of 27 subject areas with a positive relation to the ranking appears to have good results. If we consider how recently the OA phenomenon occurred, results support its goodness for multiple fields of studies. A qualitative analysis may be helpful in understanding the low success of OA in some subject areas, but that extends beyond the scope of this paper. Although descriptive statistics show prevalence in the first quartile of born OA journals, the results of regression have not confirmed a significant relationship. A final word is devoted to the fact that no significant effect has been found for the types of publishers, despite the fact that the lower number of independent scholarly publishers in comparison to other types of publishers is still low; the probability of obtaining high visibility seems to be equal.

6. Conclusions

In this paper, we have provided an integrated and novel approach to understanding the features related to the ranking of top gold OA journals. We have provided several descriptive statistics, and we have successfully applied a binary logistic regression to test all seven of the variables by their relationship with the ranking. To summarize, we have found that English is significant with a positive sign, while Multilingual and Spanish are significant with negative signs, and there are no signs of significance for other tested languages. Again, we have found that journals owned by organizations located in the USA and in the UK are positively related to the ranking. Regarding the years of existence of the journals, we have found that this variable is not related to the rankings; conversely, years in the DOAJ repository are positively related with the rankings. In other words, *coeteris paribus*, increasing the permanence in the DOAJ repository increases a journal's ranking. We have also identified the fields of study where gold OA has reached positive results in terms of ranking. Again, we have not found a significant relationship between ranking and journals launched as OA or converted and types of publishers.

Moreover, our research has brought to light an important question about funding models; because top ranking journals tend to require a fee to publish, this situation can limit *de facto* research sharing in gold OA for those who cannot support payment of publication fees.

Consider the Berlin OA declaration (2003) that states: "[. . .] *mission of disseminating knowledge is only half complete if the information is not made widely and readily available to society* [. . .]". It can be said that gold OA has reached important results for what concerns the diffusion, but it is still struggling to achieve widespread high ranking. We believe that our findings might be helpful to the OA cause, particularly for the fields of study where journals have not yet reached high rankings.

This paper has three limitations: First, it only takes into consideration journals indexed in the Scopus dataset. Second, the analysis is limited to journals indexed on 12/31/2012; a temporal evolution study could be helpful to avoid potential extemporaneous situations. Third, this research focuses only on gold OA journals; analyzing the same variables and comparing them to green OA and traditional publishing methods could offer more insight. Thus, further studies are required to improve research findings.

Acknowledgements

Authors are grateful to the two anonymous reviewers for their valuable comments and suggestions that improved this paper. Any remaining errors or shortcomings are the sole responsibility of the authors.

References

Antelman, K. (2004). Do open-access articles have a greater research impact? *College & Research Libraries*, *65*(5), 372–382.

Benos, D. J., Fabres, J., Farmer, J., Gutierrez, J. P., Hennessy, K., Kosek, D., et al. (2005). Ethics and scientific publication. *Advances in Physiology Education*, *29*(2), 59–74.
Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities. (2003). http://openaccess.mpg.de/286432/Berlin-Declaration
Björk, B. C. (2004). Open access to scientific publications: An analysis of the barriers to change? *Information Research*, *9*(2), 170.
Björk, B. C., & Solomon, D. (2012). Open access versus subscription journals: A comparison of scientific impact. *BMC Medicine*, *10*(1), 73.
Bornmann, L., Marx, W., Gasparyan, A. Y., & Kitas, G. D. (2012). Diversity, value and limitations of the journal impact factor and alternative metrics. *Rheumatology International*, *32*(7), 1861–1867.
Craig, I. D., Plume, A. M., McVeigh, M. E., 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.
Eysenbach, G. (2006). Citation advantage of open access articles. *PLoS Biology*, *4*(5), e157.
Falagas, M. E., Kouranos, V. D., Arencibia-Jorge, R., & Karageorgopoulos, D. E. (2008). Comparison of SCImago journal rank indicator with journal impact factor. *The FASEB Journal*, *22*(8), 2623–2628.
Freese, J., & Long, J. S. (2006). *Regression models for categorical dependent variables using Stata*. College Station, TX: Stata Press.
Giglia, E. (2010). The impact factor of open access journals: Data and trends. *E-Prints in Library and Information Science*,
González-Pereira, B., Guerrero-Bote, V. P., & Moya-Anegón, F. (2010). A new approach to the metric of journals' scientific prestige: The SJR indicator. *Journal of Informetrics*, *4*(3), 379–391.
Gumpenberger, C., Ovalle-Perandones, M. A., & Gorraiz, J. (2013). On the impact of gold open access journals. *Scientometrics*, *96*(1), 221–238.
Harnad, S., Brody, T., Vallières, F., Carr, L., Hitchcock, S., Gingras, Y., et al. (2004). The access/impact problem and the green and gold roads to open access. *Serials Review*, *30*(4), 310–314.
Hayes, H. (2005). *Digital repositories: Helping universities and colleges* Briefing paper: Higher education sector. JISC.
Lancho-Barrantes, B. S., Guerrero-Bote, V. P., & Moya-Anegón, F. (2010). What lies behind the averages and significance of citation indicators in different disciplines? *Journal of Information Science*, *36*(3), 371–382.
Lawrence, S. (2001). Online or invisible. *Nature*, *411*(6837), 521.
Lee, H., & Shin, J. (2014). Measuring journal performance for multidisciplinary research: An efficiency perspective. *Journal of Informetrics*, *8*(1), 77–88.
Leydesdorff, L., de Moya-Anegón, F., & Guerrero-Bote, V. P. (2010). Journal maps on the basis of Scopus data: A comparison with the journal citation reports of the ISI. *Journal of the American Society for Information Science and Technology*, *61*(2), 352–369.
Leydesdorff, L., & Opthof, T. (2010). Scopus's source normalized impact per paper (SNIP) versus a journal impact factor based on fractional counting of citations. *Journal of the American Society for Information Science and Technology*, *61*(11), 2365–2369.
Leydesdorff, L., Radicchi, F., Bornmann, L., Castellano, C., & Nooy, W. (2013). Field-normalized impact factors (IFs): A comparison of rescaling and fractionally counted IFs. *Journal of the American Society for Information Science and Technology*, *64*(11), 2299–2309.
Lobachev, S. (2008). Top languages in global information production. *Partnership: The Canadian Journal of Library and Information Practice and Research*, *3*(2).
McVeigh, M. E. (2004). *Open access journals in the ISI citation databases: Analysis of impact factors and citation patterns: A citation study from Thomson Scientific*. Thomson Scientific.
Miguel, S., Chinchilla-Rodriguez, Z., & de Moya-Anegón, F. (2011). Open access and Scopus: A new approach to scientific visibility from the standpoint of access. *Journal of the American Society for Information Science and Technology*, *62*(6), 1130–1145.
Moed, H. F. (2007). The effect of open access on citation impact: An analysis of ArXiv's condensed matter section. *Journal of the American Society for Information Science and Technology*, *58*(13), 2047–2054.
Pinfield, S. (2005). A mandate to self archive? The role of open access institutional repositories. *Serials: The Journal for the Serials Community*, *18*(1), 30–34.
Prosser, D. (2003). Institutional repositories and open access: The future of scholarly communication. *Information Services and Use*, *23*(2), 167–170.
SCImago. (2007). *SJR – SCImago Journal & Country Rank*. Retrieved from. http://www.scimagojr.com
Seglen, P. O. (1997). Why the impact factor of journals should not be used for evaluating research. *BMJ: British Medical Journal*, *314*(7079), 498.
Solomon, D. (2013). Types of open access publishers in Scopus. *Publications*, *1*(1), 16–26.
Solomon, D. J., Laakso, M., & Björk, B. C. (2013). A longitudinal comparison of citation rates and growth among open access journals. *Journal of Informetrics*, *7*(3), 642–650.