The cost of Open Access: Comparing public projects' budgets and Article Processing Charges expenditure

Patricia Alonso-Álvarez[1,2], Pablo Sastrón-Toledo[1,2], Jorge Mañana-Rodriguez[3]

[1]LEMI (Laboratory on Metric Information Studies), Department of Library and Information Sciences, Carlos III University of Madrid, Calle Madrid 126, 28903, Madrid, Spain
[2]Research Institute for Higher Education and Science (INAECU), UAM-UC3M, 28903, Getafe, Spain
[3]Spanish National Research Council, 113 Serrano Str., 28006, Madrid (Spain)

Patricia Alonso-Álvarez, patricia.alonso@uc3m.es 0000-0002-9305-6024
Pablo Sastrón-Toledo, psastron@bib.ucm3.es 0000-0003-0795-7182 (Corresponding author)
Jorge Mañana-Rodríguez, jorge.manana@csic.es 0000-0002-0717-5271

Abstract
Open Access (OA) publication often entails payment of Article Processing Charges (APCs), particularly in the so-called Hybrid and Gold journals. The growth of Gold OA publications linked to the development of OA mandates has forced funders, research institutions, and researchers to develop strategies to pay APCs. Thus, this research tries to estimate the percentage of the budget of the projects funded by the Spanish State Plan for the Generation of Knowledge and Scientific and Technological Strengthening of the R&D&I, Spain's two main public project funding calls in Spain. The period studied is 2013-2019. Additionally, we study the relationships between publication intensity, funding attraction, and the availability of OA journals with APC expenditure at the area level. The results show that €45.87 million were spent on APCs, with most projects spending 3-8% of their budgets. However, numerous outliers with rates over 10% suggest further study on the role of APCs in the financial performance of the research activity. Estimations shown in the paper have to be taken cautiously as the APCs of the publications related to the projects studied in this analysis might not have been fully paid with the projects' budget but using other funding strategies. Further research is needed to address the researchers' decisions better when paying APCs. Finally, publication intensity, funding attraction, and the availability of OA journals are highly correlated with the investment in APCs in different fields. These results show that the current APC framework affects disciplines differently and raises questions about alternative publishing and funding models.

Keywords
Article Processing Charges; Project funding; Open Access; Open Science

Introduction
Open access, Article Processing Charges, and open access mandates
The scientific literature reflects the growth of open-access (hereafter OA) publications, especially Gold OA (Björk et al., 2010; Khoo, 2019). Gold and Hybrid OA routes usually imply immediately making openly available an article preceded by the payment of Article Processing Charges (APCs). Journals with APCs comprise 30% of the OA journals but publish 56% of the research (Crawford, 2018). OA journals that do not charge APCs include the Diamond route, which is considered a subcategory within Gold journals. APC prices present great variability across fields, journals, and countries. First, fields where grant funding is expected present higher APCs, such as Science, Technology, Engineering, and Mathematics (STEM) related fields, leaving the Social Sciences and Humanities (SSH) with lower rates (Solomon, 2014). Second, the journals' impact factor (JIF hereafter) and prestige positively correlate to the APCs price (Björk & Solomon, 2015; Pollock & Michael, 2019). Third, journals in low-income countries charge the lowest APC prices against journals from the global North that present the highest rates of APCs (Solomon & Björk, 2012a).
Moreover, the rapid growth of OA and its incidence in the free online availability of research has crystallized into a series of so-called OA mandates. These mandates are normative documents

developed by policymakers and research-funding organizations. They contain regulations on the OA publication of funded research. As an example of the early adoption of OA mandates, the UK Higher Education Funding Council (HEFCE) requests the immediate deposit of peer-reviewed articles funded by the Council in institutional repositories since 2013. At the European level, Plan S was launched in September 2018 as an initiative to make publications resulting from either public or private-funded projects freely available (Plan S, 2023). Many other later examples of country-level mandates are available in the Registry of Open Access Repositories Mandatory Archiving Policies (ROARMAP, available at: https://roarmap.eprints.org/). The progressive adhesion of countries and institutions to OA mandates has made open publication a requirement rather than the result of individual authors' commitment to Open Science and Open Access.

Thus, the rise of editorial models that rely on paying APCs and the need to comply with OA mandates lead to a scenario where OA publishing is becoming more popular and extended among researchers. However, in the majority of the cases, it is inevitably linked to the affordability of the required publishing fees. Scholars have already warned about the harmful effects of the current model of OA publishing, especially in terms of increasing inequalities in the academic community (Ross-Hellauer et al., 2022). This scenario has been seen as one that 'compounds inequities' (Ross-Hellauer, 2022), where only wealthy researchers, institutions, and regions can afford to publish in OA.

Open access and research fields

To a different extent, inequalities referred to above are also present when comparing different fields of knowledge. Research fields generally present significant behavioral and organizational differences (Whitley, 2000 [1984]), which directly or indirectly might affect OA researchers' publication practices and APC funding strategies. For instance, the degree of publication intensity varies between fields (Moed, 1985). Hence, researchers from fields with higher publication rates willing –or required– to publish in OA are more likely to face APCs. Another example would be the capacity of certain research areas to attract funding, which might imply that fields with lower resources cannot afford the payment of APCs. Last, the availability of OA journals has been proven to vary across fields. In their research, Rodriguez & Guns (2022) show how 30% of the journals in microbiology –indexed in the Web of Science– are Open Access, while the percentage decreases to 5% for economics and business. Thus, researchers in areas with low OA availability are less likely to find OA journals to publish in. In monetary terms, this might imply that these areas are less likely to find OA journals that do not require APCs and, therefore, have to invest more money in APCs.

Public funding of Article Processing Charges

As seen in the previous section, the requirement to publish in OA usually implies the payment of APCs, as Gold and Hybrid journals are currently the most common venues for OA publication. To adapt to this new reality, funding agencies, research institutions, and researchers have developed different strategies to cover APCs fees.

Research funding agencies started covering APC fees in the early 2000s (Tananbaum, 2003). The assumption of publication costs by funding agencies has a crucial precedent in the European Commission Pilot initiative, which allowed researchers to fund APCs associated with finished FP7 projects (FP7 Post-Grant Open Access Pilot). In some cases (OpenAIRE), APC prices were jointly negotiated with publishers. Inspired by European initiatives, APCs are eligible costs of the budgets of Spanish funding grants. Ferrer-Sapena et al. (2021) estimated the cost of APCs in four fields of the Social Sciences (Humanities, Sociology, Information & Library Science, and Education & Educational Research & Communication) between 2012 and 2019 in €481,120.4. They also found that 53% of the articles in their dataset acknowledged governmental funds from

either national or regional funders; the European Union and universities were acknowledged in 20-25% of the articles.

Research institutions have also developed their own strategies to accommodate the payment of APCs in their budgets. Mainly, they have set central funds managed by their libraries to cover researchers' APCs as well as transformative agreements with publishers (Borrego, 2023). Transformative agreements are contracts between publishers and institutions where the latter pay an amount of money to cover the fees to publish in OA for a limited number of articles. However, they might also include subscription costs or, in some cases, unlimited access to publications in the journals covered by the contract (Borrego et al., 2021). Several studies have sought to quantify the amounts of public investment transferred to publishers as APCs covering different regions or time periods. The findings of Aasheim et al. (2019) indicate that the APCs paid by European institutions have nearly doubled from 2005 to 2018. Quadery et al. (2019) also estimate that a redirection of €150 million to the journals publishing publicly funded research would be necessary to accomplish Plan S, which is a cost that will have to be ultimately assumed by research funders. A recent study by Butler et al. (2023) estimates the APCs paid to journals belonging to Elsevier, Sage, Springer-Nature, Taylor & Francis, and Wiley between 2015 and 2018. Their analysis estimates the APC expenditure in journals belonging to the stated publishers to be $1,06 billion.

Qualitative research based on surveys has found that scholars also use their personal funds to cope with APCs, especially in specific research areas and circumstances. Solomon (2014) reflects that using personal funds is more common in SSH fields, while Solomon and Björk (2012) find that using personal funds is more frequent in low-income countries. Still, some research indicates that grant funding is generally the preferred method to finance APCs (Halevi & Walsh, 2021; Swan & Brown, 2004). This fact raises both financial and ethical questions. First, about the amount of public money spent in APCs through projects' budgets. Second, concerns about whether or not it is appropriate that projects' budgets are used to pay for publication costs rather than for scientific needs, such as buying equipment or hiring researchers or technical staff.

Project funding and APCs in Spain

Competitive projects are the primary funding source for researchers in Spain and are highly valuable for academic promotion (Molas-Gallart, 2012). Scholars have estimated that around 37% of the funds dedicated to scientific research in Spain are allocated via project funding (Fernández-Zubieta & Ramos-Vielba, 2017). Zacharewicz et al. (2019) concluded that Spain is the eighth European country with a higher share of scientific investment allocated via project funding. This emphasis on project funding has placed projects as a key instrument for basic funding in Spain.

The central research funding organization that allocates funding from the State Secretary of Research is the Spanish Research State Agency (AEI). The two main AEI's calls for basic research funding are Challenges of Society (oriented research toward pre-existing goals) and Generation of Knowledge (non-oriented research), which comprise between 37% and 56% of the total budget of AEI. Both calls (hereafter CKG - Challenges and Knowledge Generation) are granted to research groups –rather than individual researchers– and include APCs as eligible costs in their budget schemes. To do so, the CKG calls refer to the Spanish Law of Science, Technology, and Innovation. As our study only includes the period between 2013 and 2019, the calls refer to the 2011 Law (14/2011) rather than to its modification in 2022, which refines and extends the OA mandate. The Law 14/2011 includes a specific mandate stating that "researchers whose research activity is mainly financed with funds from the General State Budget shall make public a digital version of the final version of the contents accepted for publication in serial or periodical research publications, as soon as possible, but no later than twelve months after the official date of publication" (España, 2011, art. 37). However, it allows the transference of

industrial and intellectual rights to third upon the authors' request and when research results have to be protected (España, 2011). In practical terms, the exception allows the authors to skip the OA mandate if they publish in a journal whose license prevents them from openly publishing the results. Furthermore, there are no consequences for non-compliance with Spanish regulations. Previous research in Spain has analyzed the expenditure in APCs by research funding or fields (Ferrer-Sapena et al., 2021) or regions. For instance, a recent study by Consorcio Madroño (2020) –a network of 6 universities from the Madrid Autonomous Region– estimated that €1,926,643 was spent on the coverage of APCs in 2019. However, there has been no effort to analyze public projects' expenditure on APCs.

Objectives

This study focuses on the CKG call to calculate the amount spent in APCs from the projects' budgets. However, since we lack information on the APCs' paying behavior of researchers funded by CKG calls, we can only estimate the APCs' fees paid per the publications associated with CKG projects and compare them with the project's budget. As specified above, researchers might use different strategies to deliver APCs, even if their research is funded by competitive funds. Therefore, our analysis does not establish a direct relationship between the money spent on APCs and the projects' budget. Instead, the study gives a general picture of the expenditure in APCs within the context of the projects' budgets, aiming to answer two main research questions:

> RQ1: *What is the amount of money spent in the APCs of publications associated with projects of the Spanish CKG call?*
>
> RQ2: *What percentage of the project budget is equivalent to the APC expenditure resulting from project publications?*

In addition, we aim to study to what extent research areas influence APC expenditure. Based on the area differences expressed in the previous section. As OA mandates operate in CGK calls, most of the publications of the CGK projects should be openly available. However, projects with higher publication rates might also explore other strategies to fulfill the mandate, such as publishing their research outputs in repositories or non-APC journals. Also, as some studies have shown (Solomon, 2014), resource-intensive fields where funding is expected (mostly STEM) present higher APCs. Thus, we expected the total expenditure in APCs to be higher. Finally, we want to explore if the availability of journals in OA may influence the current spending in APCs. Therefore, we propose the following complementary questions:

> RQ3*: Do projects that publish more academic articles pay more money in APCs?*
>
> RQ4*: Do projects with higher budgets pay more money in APCs?*
>
> RQ5: *Is there a correlation between the expenditure in APC and the availability of OA journals along the different fields of knowledge?*

Data

CKG projects data

We constructed a dataset containing all the information on the projects accepted from 2013 to 2019. CKG projects were obtained by downloading the project award resolutions on the Spanish National Research Agency's website. These resolutions are in PDF format; therefore, the data was previously structured using the Python library tabula.py. Finally, information was obtained

on 19,766 projects from 2013 to 2019, a total of €2,319 million financed. The dataset contains information related to the projects, such as funding, institution, research area, and grant number.

Publications dataset

Articles from CKG projects were retrieved from Clarivate Analytics Web of Science (WoS) - 'Core Collection' and Science, Social Sciences and Arts & Humanities Citation Indexes (SCI, SSCI, AHCI, and ESCI) databases - using the funding acknowledgments field. This field was cross-referenced with the CKG grant numbers, thus obtaining the scientific production of each project awarded. We retrieved 105,559 publications from 2013-2019 (collected on 27 October 2021), containing all the articles that acknowledged the funded projects.

A total of 66,980 publications were published in OA (63,4% of the total number of publications produced by the projects). Half were published in Gold journals (25,495) and Hybrid journals (5,890). It should be noted that the category Gold is taken from the Web of Science, so it includes all journals that are fully open, including journals with no APC –i.e., Diamond.

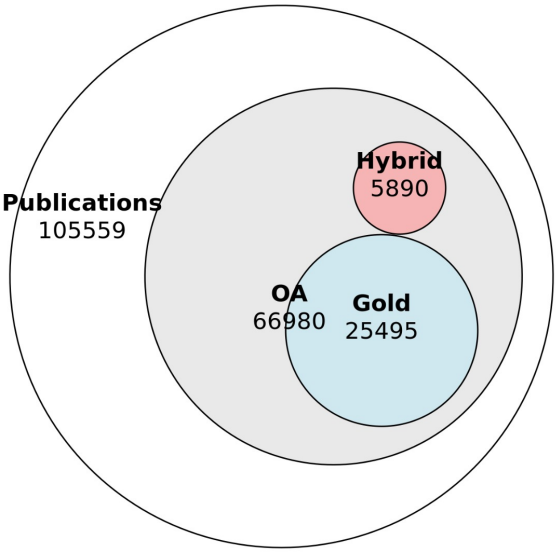


Figure 1. Classification of publications regarding open-access

APC database

The APC database used in this analysis comes from four different sources:

- The first dataset contains APC information registered in the OpenAPC initiative (https://openapc.net/), a non-profit platform aggregating voluntary data from research institutions. Open APC offers data on the cost of all APCs per article paid by the collaborating partners. APCs for the same journal and year vary depending on the reporter agency, so we calculated the minimum APC paid per journal per year. We used it for the analysis so costs are not overestimated. The dataset contained a combination of 29,710 journal-years.
- The second dataset (Matthias, 2020) contains the APC cost of ten publishers (Cambridge, Copernicus, Elsevier, Hindawi, Nature, Oup, Sage, Springer-Nature, Taylor & Francis, and Wiley). This information was retrieved from the publishers' web pages through the Internet Archive's Wayback Machine, thus reflecting the different prices of APCs over time. For more information on the APCs retrieval process, refer to Matthias (2020). Once the data was processed, the dataset contained a combination of 65,803 journal-years.
- The third dataset (Morrison, 2021) contains mainly APC listed on the Directory of Open Access Journals (DOAJ, https://doaj.org/). Butler et al. (2023) also point out that the

information of this dataset includes "a previous dataset from DOAJ, Crawford (2019), DOAJ, OpenAPC and Morrison (2019), including frequent manual checks on journal websites". For more information on the APCs retrieval process, refer to Morrison (2021). Once the data was processed, the dataset contained a combination of 102,860 journal-years.

- Finally, the data was refined using Butler et al.'s (2022) dataset, which included some new journals and a refined version of Matthias (2020) and Morrison (2021). The update of Butler et al.'s (2022) dataset happened in the process of creating this work. The dataset contained a combination of 18,797 journal-years.

A criterion was established to cross-reference the publications with their APCs cost in an iterative process. We first used the OpenAPC database since it is the dataset that matches the highest number of publications, and the APCs are registered in the same currency as the project funding (euros). Then, Matthias (2020), Morrison (2020) and Butler et al. (2022) APCs were converted to euros when needed (taking into account the value of the year of the given journal) and cross-referenced with the remaining records. In the end, 29,762 publications had an assigned APC from the 31,484 total of Gold and Hybrid publications (94% hit rate). Therefore, the results underestimate the actual expense of APC in the studied period. Unmatched publications journals' did not figure in the APC database. We argue that one of the reasons underlying the absence of some of the journals could be due to the limited regional or language coverage of the databases. Specifically, we observe 37% of the unmatched publications being published in hispanic journals (Spain and Latin American journals), which might be under-represented in the used databases. Other absent journals might be due to the compilation processes of the used datasets –for instance, Open APC relies on self-reporting from universities, and Matthias (2020) focuses on journals from big publishers. Future research may focus on improving this ratio and include the data that has not matched, thus improving the existing databases. A version of the compilation of used databases can be found in Alonso-Álvarez et al. (2024)[1].

OAL dataset

The OAL dataset comprised a subset of the publications' dataset containing information about 1,607 articles published between 2019 and 2021, with funding acknowledgments from 2019 projects. This subset allows us to test if there is a linear relationship between the expenditure in APC (as a fraction of the total funding) and the availability of OA journals in the various fields of knowledge calculating the OAL indicator (Mañana-Rodríguez & Guns, 2022).

Methods

Information processing and visualization

The software used to process the data was Python 3.9, mainly through the Pandas library (The Pandas development team, 2023). Tableau 2021.3 was used for the visualization in Figure 1.

Descriptive statistics and visualization

To answer RQ1, RQ2, RQ3, and RQ4, we use descriptive statistics to summarize and characterize the data, while simple correlation analyses are performed to discern potential associations between variables. To illustrate the relationship between APC expenditure and project funding, we also rely on a box-plot visualization.

OAL calculation

[1] Our work was merely compilation work. In the case of using the same data we highly recommend to refer to the original datasets.

For the last question (RQ5), we used the OAL indicator (Mañana-Rodríguez & Guns, 2022). The indicator provides a proxy for the starting position of projects in different fields concerning OA publication mandates: less abundance of OA journals without APCs in the fields in which the projects' published results might affect the proportion of funding used for APC expenses. The requirement of OA publication is constant for projects in all fields, but the proportion of OA journals and journals with and without APCs is not. Suppose all publications acknowledging funding from the projects of a given field of knowledge are published within a single WoS subject category. In that case, it is possible to directly quantify the availability of OA journals and journals with and without APC and compare them across projects in a given field of knowledge. Nevertheless, the articles acknowledging funding from projects in any field of knowledge are scattered through several subject categories. OAL is intended to provide a summary of the abundance of OA and APC journals considering two factors: a) the distribution of articles (for a given set of projects within a subject category) across subject categories and b) the proportion of journals publishing in OA or with APCs in each of those subject categories.
The likelihood for a given project to be published in an open-access journal is then the average of disciplinary open-access likelihoods, weighted by the project's share of publications in each discipline, expressed as a percentage (following Mañana-Rodríguez & Guns, 2022):

$$OAL = \sum_{i=1}^{n} \left(\frac{P_i}{\sum_k P_k} \frac{OAJ_i}{J_i} \right)$$

Where:

n: is the number of disciplines

i $(i = 1, \dots, n)$: is a given discipline

P_i: is the number of articles published in the field i
OAJ_i: is the number of open-access journals in a discipline i
J_i: is the total number of journals in a discipline i

The indicator expresses the likelihood (ceteris paribus) of the research projects in a given field of knowledge to be published in OA, given the availability of OA journals (and journals with APCs). The indicator is a weighted average of the likelihoods of OA journals (and journals with APCs) per discipline, weighted by the share of publications in each discipline for the projects in each field of knowledge (the disciplines of the projects are those given by the funding organization, consisting of 17 broad fields).
As a limitation of the indicator, it is relevant to mention that the different volumes of articles published by journals, among other factors mediating the choice of journals, are not included in its calculation.
Since the publication of articles stemming from funded projects can take several years, we decided to choose, as a proxy for the values of OAL, the articles associated with projects funded in 2019 published in the 2019-2021 period (1,607 articles). Then, we matched the ISSN or E-ISSN of the journals with the Essential Science Indicators (ESI hereafter) list of journals (using the 2019 list, the same year the projects were awarded) to retrieve their research fields. ESI contains 11,855 journals classified into 22 research fields, and each journal is classified into only one of them. We found 90 journals for which no ESI research field was retrieved; this can be explained mainly by two factors: on the one hand, some journals belong to disciplines in the humanities, for which there is no equivalent research field in ESI and, on the other hand, some journals in the 2020 and 2021 sets were not present in ESI 2019's master list. Since the objective of the OAL analysis is to serve only as a proxy, we anticipate that the error will not significantly

affect the results. Along this step, we identified OA journals and journals with APCs in the ESI master list using DOAJ (Directory of Open Access Journals).
In the next step, we calculated the OAL indicator using the 17 research areas established by the AEI –areas ANEP of the Spanish Ministry of Science and Innovation (https://www.aei.gob.es/areas-tematicas/areas-tematicas)– as the unit of analysis. For the calculation of the indicator, we used the percentages of OA journals by ESI field, the % of journals with APC, and the percentages of articles published in journals belonging to each ESI field for each field of knowledge (for all projects in each of the 17 fields of knowledge).
Finally, we calculated the average percentage of the total budget spent in APC by the projects of each of the 17 areas ANEP and its correlations with the OAL indicator for OA journals and the OAL indicator for journals with APCs.

Results

Project descriptives

The table below (Table 1) shows the main descriptives of the projects analyzed by research area. Research areas are based on the categorization of areas ANEP. The table shows the number awarded by research area between 2013 and 2019 and the significant differences between areas, which are also observable in terms of the project's average budget and average publications per project. These differences, especially in the budgets and publications, are expected given each research field's different needs and behaviors (Moed, 1985; Whitley, 2000 [1984]). On average, Biomedicine and Life Sciences projects receive higher budgets than other areas, especially those related to Social Sciences, Economy, and Law. Regarding publications per project, some areas, such as Mathematics or Physics, show a high productivity while others publish fewer articles. Generally, higher-budget projects correlate with more publications (Pearson correlation coefficient 0.42). This phenomenon might also be explained by the different publication habits of each discipline and the limitation of our analysis to a single documental type: academic articles published in a restricted database, Web of Science.

Table 1. Average project's budget in euros per research area

Research Area	*Number of projects per area*	*Average project's budget (€)*	*Average publications per project*
Environmental Science	1,740	145,446	6
Biomedicine	1,614	206,185	4
ICT	1,598	116,987	8
Life Science and Biotechnology	1,516	203,689	4
Agriculture and Food-Science	1,486	149,498	5
Physics	1,414	192,051	9
Chemical Science	1,251	139,807	11
History and Art*	1,107	43,233	1
Materials Science	1,023	123,336	10
Philology and Philosophy*	1,004	39,847	1
Industrial and Civil Engineering	1,002	121,732	5
Law	843	26,595	0.2
Economics	722	33,084	6
Social Sciences	715	48,999	2
Mathematics	696	53,640	13
Psychology	680	67,403	4
Energy and Transport	604	134,846	6
Education	409	40,964	2
History and Archeology*	110	40,777	0.4
Culture: Philology, Literature and Arts*	106	39,867	0.3
Language, Mind and Thought*	64	40,686	1

*Both "History and Art" and "Philosophy and Philology" were replaced in 2018 by the areas "Language, mind and thought," "History and Archeology," and "Culture: Philology, literature, and arts." Both denominations are presented because they are not strictly equivalent, but note that the last ones contain data from 2018 onwards.

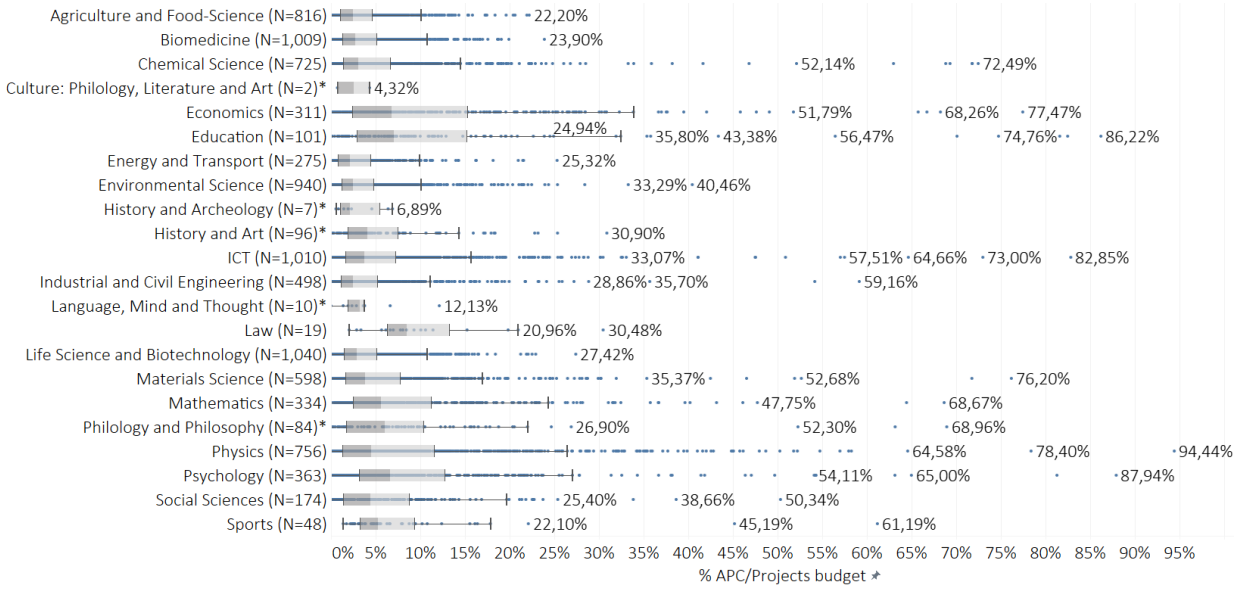
Total APC expenditure

The analysis reveals that 29,762 publications were published in 1,093 Gold journals and 1,513 Hybrid journals. The publication of these articles cost €45.87 million in APCs from 2013 to 2019: €32.22 million in Gold journals and €13.65 million in Hybrid journals. This amount equals 1.98% of the total budget financed by the CKG programs from 2013 to 2019. It is worth highlighting that the results underestimate the actual expense of APC since these publications represent 94% of the total Gold and Hybrid publications identified.

APC expenditure in the sample of projects fully covered by the APCs database

We further analyzed a sample of projects whose Gold and Hybrid publications (29,762 publications) were fully covered by our database (8,769 projects, 44% of the total projects). Limiting our analysis to fully covered projects, we are able to get a more accurate picture of the relationship between APC expenses and projects' budgets. Figure 2 illustrates the APCs costs in relation to the projects' budget by research area. For the estimation, we calculated the percentage of the projects' funds that would have been spent in APCs had all the APCs been paid with the projects' budget. However, it is important to note that the numbers in Figure 2 do not present a direct relationship, as APCs might have been paid using one of the other strategies specified in the introductory section, such as transformative agreements or personal funds.

Figure 2 shows that the median of the share of the projects' budget compared to the expenditure in APCs tends to be between 3-8% of the funds. However, some areas present higher values than others. For instance, in the case of Law, the higher average value is biased by the small amount of money available for the projects (see Table 1). Also, when looking at the outliers, there are some projects where the amount of APC exceeded half of the project's budget. These high ratios are usually due to the small amount of project funding, the large number of articles published, or the high prices of the APCs in some journals. For instance, the Economics project that spent the equivalent of 77.47% of its budget in APCs had a total budget of €9,559 and produced two articles with APCs of €4,455 and €2,950. Also, in the case of Physics, an area with many outliers, the average number of publications per project is higher than in other areas (9 publications), so for those projects with low budgets (i.e., theoretical physics), the amount of money associated with APCs is relatively high when compared to the total budget of the projects.



*Both "History and Art" and "Philosophy and Philology" were replaced in 2018 by the areas "Language, Mind and Thought," "History and Archeology" and "Culture: Philology, Literature, and Arts." Both denominations are presented because they are not strictly equivalent, but note that the last ones contain data from 2018 onwards.

Figure 2. Percentage of projects' budget estimated to be spent in APC by ANEP areas.

Research area differences

Publication intensity and funding attraction capacity and their relation to APCs expenditure

When exploring the possible effects of research areas on APCs expenditure, we analyzed the correlations between both the amount of articles published per project (RQ3) and the projects' budget (RQ4) and the APCs expenditure. To answer these questions, we used the same sample as in the previous section –which contained only the projects whose publications with APCs were fully covered by our database. We found the two correlations to be positive. The correlation coefficient between the number of articles published per project and the expenditure in APCs was 0.56, showing that research areas with higher publication intensity spend more money on processing charges. Regarding projects' budgets, the correlation coefficient was 0.97, highlighting the increased investment in APCs of projects with more resources. In sum, we note that fields with higher publication rates and high funding attraction capacity are likelier to spend more on APCs when deciding to publish their results in OA.

Open Access journals availability

One of the reasons that might contribute to explaining the different percentages spent in APCs by research area is the availability of open-access journals. A possible hypothesis is that areas with a lower share of OA journals are less likely to find OA journals that do not require the payment of an APC (Mañana-Rodriguez & Guns, 2022). Table 2 shows the OAL calculation for all OA journals, journals with APCs, and the average project budget presumably spent on APCs for 2019. The OAL indicator has been rescaled into a 0-100 scale in Table 2 in order to facilitate the presentation of small values. As stated before, we used 2019 for the OAL calculation as it is the most recent year in our dataset. The OAL indicator is interpreted as the likelihood for a given project to publish in an open-access journal –or a journal with APC–, given the project's set of publications and the number of open-access journals –or journals with APC– in a discipline.

The table shows that the OAL varies strongly between disciplines, from values close to 20 in Life Sciences and Environmental Sciences to values below 10 in disciplines related to SSH fields (except for Mathematics). Regarding journals with APC, they represent on average 64% of the total OA journals. However, these values vary between fields. Table 2 shows that, in general, the OAL for journals with APCs follows a similar pattern, although some disciplines have different values than expected. For instance, Environmental Sciences shows a higher proportion of non-APC journals than its counterparts. Finally, the fraction of expenditure in APCs in relation to the projects' budget also reflects clear differences between fields. Fields with higher rates are commonly from the Social Sciences and Humanities, with some exceptions (Physics, and Energy and Transport).

Table 2. OAL, OAL over APC, % of expenditure on APCs, and average APC per research field.

Field of knowledge of the projects	OAL (OA journals)	OAL (OA journals with APC)	APC / avg. project budget
Life Science and Biotechnology	23.07	17.72	1.31%
Agriculture and Food-Science	22.61	16.15	0.89%
Environmental Science	20.88	13.68	1.62%
Biomedicine	20.73	15.56	1.43%

Industrial and Civil Engineering	15.98	10.54	1.31%
History and Art	15.92	9.98	3.31%
Materials Science	15.57	9.58	2.23%
Physics	15.46	9.92	2.71%
Chemical Science	14.92	9.08	3.25%
ICT	12.65	8.11	2.41%
Psychology	11.67	8.47	2.81%
Energy and Transport	10.80	6.57	3.64%
Social Sciences	10.46	5.35	4.91%
Philosophy and Philology	9.55	6.50	6.43%
Education	9.38	5.48	6.42%
Mathematics	8.89	4.94	2.92%
Economics	8.46	4.39	5.51%

The areas "Language, Mind and Thought","History and Archeology" and "Culture: Philology, Literature and Arts" were excluded from the analysis because there is no equivalent research field in ESI for these fields (Humanities).

Table 3 shows the correlations between the OAL indicator – both for the OA journals and the journals with APC– and the expenditure in APCs in relation to the project budget (the third column in Table 2). The results show a moderate, negative correlation between the availability of open-access journals in a research area and the amount of money spent on APCs associated with a CKG project. Thus, researchers in fields with lower availability of OA journals have higher expenditures in APCs and vice versa. We consider this result to be merely exploratory, possibly guiding future research. Moreover, our sample is limited only to Spain and journals indexed in the Web of Science, which may induce biases, such as undermining the representativeness of SSH fields.

Table 3. Correlations between OAL and % APC

ρ[OAL (% of OA journals); %APC]	ρ[OAL (OAL (% of journals with APC (Over total); %APC]
-0.80	-0.78

Discussion and conclusion

In Spanish publicly funded projects, the costs of OA publishing are considered eligible under the projects' budget. As OA publishing grows, researchers and, therefore, research agencies have to spend more money on paying APCs. Thus, the first research question of this study (RQ1) aimed to quantify the amount of money spent on the APCs of publications associated with projects of

the Spanish CKG call. We found that at least €45.87 million were transferred to different publishers to pay for APCs in projects awarded between 2013 and 2019. However, the number is expected to be higher as our dataset does not cover all journals where CKG research is published – 94% of Gold and Hybrid journals are covered. The amount of money paid in APCs is practically equivalent to the total budget of the CKG calls for the area of psychology in the period studied (€45.83 million). The comparison makes it inevitable to reflect on alternative uses of this money, both related to pure scientific –such as personnel hiring or equipment– or editorial purposes, such as the development of alternative publishing models. As a result of the different OA initiatives leading to the open dissemination of research results as well as to the growth of the Gold OA model, we expect that spending in APC might increase over time. However, not only researchers are changing and adapting their dissemination practiques to current editorial models. The development of alternative publishing models is emerging as a response. As an example, the so-called Diamond journals (institutionally funded journals that do not charge fees for publishing or reading) are being subject of study, both in Europe (Arasteh, & Blake, 2024; Laakso & Multas, 2023) and worldwide (Khanna et al., 2022).

The following steps of this project will try to obtain the costs of APCs for the remaining articles, which we believe mainly belong to Spanish journals, to get a more accurate perspective of APCs expenses in Spain. To accomplish this objective, the authors of this article would like to use this space to encourage institutions to contribute to the OpenAPC database and researchers to continue expanding the coverage of existing datasets.

Aiming to contextualize APC expenses in further detail –especially in terms of what we called 'pure scientific work' in the paragraph above–, we compared them with project budgets. This comparison allowed us to understand how much this cost meant in terms of scientific funding. Thus, our second research question (RQ2) was directed to calculate the percentage of projects' budgets that equal the APC expenditure resulting from projects' publications. Using a sample of projects fully covered by our APC database, we found that the average is between 3 and 8% of the total funding of the projects. However, the numbers differ between research areas, and some disciplines show higher percentages. We showed that some projects spend the equivalent of more than half of their budget on APCs. Although we cannot confirm that all the APCs paid by one specific project come from the project's budget, these high results might contribute to the discussion on how project funding should be spent and the compatibility of fulfilling the objectives of the project while meeting the requirements of the OA mandates. Alternative strategies to cope with APCs might also be explored, as institutions and researchers might be relying on transformative agreements (Borrego, 2023) or personal funds (Solomon, 2014).

In this sense, publishing in OA might be restricted to researchers with enough funds to pay APCs. Previous research has warned about the inequities that the APC model may involve (Klebel & Ross-Hellauer, 2023), especially for countries in the Global South. Therefore, national case studies, such as this one, may contribute to the literature on the harmful effects fostered by the APC model. An interesting approach that might be explored in future research would be to compare APC expenses and allocated funding from a cross-country perspective. This analysis might contribute to studying the performance and strategies of different countries on APC payment and the effects of the model in countries with lower economic resources.

The results also raise questions about the role of CKG funding in the research activity and the researchers' behavior regarding project acknowledgment and APC payment. From an accountability point of view, it is essential to know if the publications that acknowledge CKG projects are related to the project's purpose and objectives. If public money is being spent on APCs, it is imperative to know that the publications are related to the topic of the projects. However, these issues have to be addressed from qualitative or mixed methods perspectives that help to develop strategies and mechanisms to improve the tracking process of scientific outputs. Future research could also consider developing surveys to learn more qualitatively about

researchers' habits when publishing in OA, also considering field specificities. An additional issue results from the changing funding structure across countries, which necessarily affects how researchers cover APCs. The shift in recent decades from block funding models toward competitive funding instruments has led to increased competition for resources (Lepori et al., 2023). Although Spain is not an exception to this tendency, and now most funds are allocated via competitive projects, the share of project funding has not changed significantly over the years (Reale, 2017). The enforcement of OA mandates and the rise in the prices of APCs may be detrimental to researchers who do not have access to either project or institutional funding.

A limitation regarding RQ2 relies on the relation between the project acknowledged in each publication and the strategies followed by the authors to pay the associated APCs. Even if a publication acknowledges a specific project, it does not mean that the APC has been paid with the project budget, as there are other ways to finance these costs (i.e., transformative agreements, budgets from other projects, institutional funds). However, in absolute terms, these results can bring some light to the discussion about the benefits of dominant editorial models nowadays, especially considering the ongoing discussion on double dipping (Pinfield et al., 2015; Asai, 2023). A substantial addition to this contribution would be expanding the publication dataset, but especially including in the analysis the numbers behind subscription contracts and transformative agreements to unveil the complete picture behind the financial model of academic publishers.

Regarding field differences, the analysis confirms that projects with higher publication intensity spend more money in APCs (RQ3). Although this relationship might initially sound intuitive, research areas with higher publication rates might have also developed alternative publication strategies that do not require APCs. For instance, OA mandates usually allow authors to openly publish the final version of their work after the journals' embargo period. Therefore, authors can decide whether or not they want to pay the corresponding APC or wait for their work to be open, usually with the possibility of sharing a preprint version immediately. Previous works have also found that many journals in fields such as Humanities and the Social Sciences, usually with lower publication intensity and thus small publishers, do not charge APCs, which might also affect the amount spent in APCs by these areas (Frantsvåg and Strømme, 2019). Other variables beyond the scope of our analysis could also affect the willingness to publish in OA (i.e., researchers' institutional support to publish in OA, degree of awareness of the authors to publish in OA).

The correlation analysis also confirmed that projects with higher budgets usually pay more money in APCs (RQ4). One interpretation might be that in fields where research funds are more readily available, the pressure to lower APCs is less or does not exist (Solomon, 2014). However, the absence of pressure to lower APCs might affect researchers in the same field who have less access to economic resources –ex. early career researchers, researchers from low-income countries, or, even within the same country, from regions where less funds are awarded. Moreover, research fields and institutions might also be playing a role, as institutions' specialization may affect the projects that are granted and, therefore, their publication practices (Rodríguez & Guns, 2022). However, as discussed before, the results also show some projects with lower funds whose expenditure in APCs represents a high percentage of their budget. To fully evaluate these cases, knowing more about APC paying practices would be fundamental.

Since journals themselves are important elements of this analysis, we lastly explored if there is a correlation between the expenditure in APC and the availability of OA journals along the different fields of knowledge (RQ5). Previous literature has shown the high differences in the availability of OA journals depending on the research field (Rodríguez & Guns, 2022). Regarding the OAL calculation, the correlation between the indicator and the average percentage expenditure in APCs proves to be negative. This is, on average, a higher availability of OA journals is related to a lower expenditure on APCs. Further research might clarify the relationship between the variables. In the case of the OAL for OA journals, a higher availability of OA journals might imply a higher presence of journals without APCs (Diamond journals). In the case

of the OAL for journals with APCs, an increase in the offer of Gold and Hybrid journals might force publishers to develop market strategies that lower APC prices. It must be noted that the correlations expressed here do not imply causality, nor is the objective of this paper to establish causal connections. For instance, this relationship might be mediated by factors such as the presence of projects with small budgets, hence increasing the percentage spent in APCs when publishing in research areas with low proportions of OA journals. However, it shows that the budgetary pressure for OA publishing is unequal for all fields of knowledge. If, as proposed above, we assume that money spent in APCs is detracted from other budget items directly related to pure scientific tasks, such inequalities should be considered in the design of the funding mechanism. Further research might clarify the hypothetical causal relationship between the variables, such as the presence of journals without APCs (Diamond journals) or the existence of market strategies when more journals turn to APCs models.

Finally, we acknowledged the limitation of using Web of Science data since some fields, mostly SSH fields, are underrepresented (Mongeon & Paul-Hus, 2016). However, the Spanish research evaluation system is strongly metric-based –mainly relying on bibliometric indicators–, driving researchers to prioritize publications in journals indexed in Web of Science and Scopus with high impact factors (Molas-Gallart, 2012). Therefore, our bias might not be as significant as expected. However, further research would benefit from including publications outside the Web of Science.

Acknowledgments

The authors would like to acknowledge the comments and suggestions of one anonymous reviewer on previous versions of this manuscript.

The paper is a substantially extended version of Sastron-Toledo et al. (2023), a conference proceeding presented at the 19th International Conference on Scientometrics & Informetrics ISSI 2023, July 2-5, 2023, in Bloomington, Indiana, USA.

This study was partially funded by the DOSSUET project ("Diagnosis of open science in the Spanish university system and mechanisms for its transformation and improvement") grant number PID2019-104052RB-C21. This project also holds an FPI predoctoral training contract PRE2020-092917 of the Spanish Ministry of Science and Innovation and PIPF contract from the Madrid Education, Science and Universities Office (reference: PIPF-2022/PH-HUM-25403).

We finally want to thank Leigh-Ann Butler, Lisa Matthias, Marc-André Simard, Philippe Mongeon, Stefanie Haustein, Heather Morrison, and OpenAPC for generously sharing their APC databases. This project would not have been possible without their work.

Conflict of interest

The authors have no conflicts of interest to declare relevant to this article's content.

References

Aasheim, J.H., Ahlborn, B., Ambler, C., Andrae, M., Apel, J., Becker, H.-G.,Young, P. (2019). *Datasets on fee-based open access publishing across German institutions*. Bielefeld University. https://doi.org/10.4119/UNIBI/UB.2014.18

Alonso-Álvarez, P., Sastrón-Toledo, P., & Mañana-Rodríguez, J. (2024, February 12). The cost of Open Access: Comparing public projects' budgets and Article Processing Charges expenditure [Data set]. https://doi.org/10.17605/OSF.IO/XTSE5

Arasteh, S., & Blake, O. (2024). The European landscape of institutional publishing - A synopsis of results from the DIAMAS survey. Zenodo. https://doi.org/10.5281/zenodo.10551710

Asai, S. (2023). Does double dipping occur? The case of Wiley's hybrid journals. *Scientometrics*, 128(9), 5159-5168.

Björk, B. C., Welling, P., Laakso, M., Majlender, P., Hedlund, T., & Guðnason, G. (2010). Open access to the scientific journal literature: situation 2009. *PloS one*, *5*(6), e11273.

Björk, B.C., & Solomon, D. (2015). Article processing charges in OA journals: the relationship between price and quality. *Scientometrics*, 103(2), 373–385. https://doi.org/10.1007/s11192-015-1556.

Borrego, Á., Anglada, L., & Abadal, E. (2021). Transformative agreements: Do they pave the way to open access? *Learned Publishing*, 34(2), 216–232. https://doi.org/10.1002/leap.1347

Borrego, Á. (2023). Article processing charges for open access journal publishing: A review. *Learned Publishin*g, 36(3), 359–378. https://doi.org/10.1002/leap.1558

Butler, L., Matthias, L., Simard, M-A., Mongeon, P., & Haustein, S. (2023). The Oligopoly's Shift to Open Access. How the Big Five Academic Publishers Profit from Article Processing Charges. *Quantitative Science Studies*, Advance publication. https://doi.org/10.1162/qss_a_00272

Butler, L.-A., Matthias, L., Simard, M.-A., Mongeon, P., & Haustein, S. (2022). Annual Article Processing Charges (APCs) and number of gold and hybrid open access articles in Web of Science indexed journals published by Elsevier, Sage, Springer-Nature, Taylor & Francis and Wiley 2015-2018 (Versión v1) [Data set]. Zenodo. https://doi.org/10.5281/zenodo.7086420

Consorcio Madroño (2020). *Estimación de costes de publicación en Acceso Abierto, por pago de APC (Article Processing Charges), asociados al Consorcio Madroño durante los años 2018 y 2019*. Available at https://e-archivo.uc3m.es/handle/10016/31213#preview

Crawford, W. (2018). GOAJ3: *Gold open access journals 2012–2017*. Livermore, California: Cites & Insights Books

Crawford, W. (2019). Gold Open Access 2013-2018: Articles in Journals (GOA4). Livermore, CA: Cites & Insights Books. Retrieved from https://waltcrawford.name/goa4.pdf

España (2011). *Ley 14/2011, de 1 de junio, de la Ciencia, la Tecnología y la Innovación*.

Fernández-Zubieta, A., & Ramos-Vielba, I. (2017). Research & Innovation in Spain 2016. Working Paper Series, 2017–02.

Ferrer-Sapena, Antonia; Vidal-Cabo, Christian; Aleixandre-Benavent, Rafael; Valderrama-Zurián, Juan Carlos (2021). Coste de la publicación en abierto de artículos de autoría española en cinco áreas de las ciencias sociales. *Arbor*, 197(799): a590. https://doi.org/10.3989/arbor.2021.799005

Frantsvåg, J. E., & Strømme, T. E. (2019). Few open access journals are compliant with Plan S. Publications, 7(2), 26. https://doi.org/10.3390/publications7020026

Halevi, G., & Walsh, S. (2021). Faculty Attitudes Towards Article Processing Charges for Open Access Articles. *Publishing Research Quarterly*, 37(3), 384–398. https://doi.org/10.1007/s12109-021-09820-x

Khanna, S., Ball, J., Alperin, J. P., & Willinsky, J. (2022). Recalibrating the scope of scholarly publishing: A modest step in a vast decolonization process. Quantitative Science Studies, 1-19. https://doi.org/10.1162/qss_a_00228

Khoo, S. (2019). Article processing charge hyperinflation and price insensitivity: An open access sequel to the serials crisis. *Liber Quarterly*, 29(1).

Klebel, T., & Ross-Hellauer, T. (2023). The APC-barrier and its effect on stratification in open access publishing. *Quantitative Science Studies*, 4(1), 22–43. https://doi.org/10.1162/qss_a_00245

Laakso, M., & Multas, A. M. (2023). European scholarly journals from small-and mid-size publishers: mapping journals and public funding mechanisms. Science and Public Policy, 50(3), 445-456. https://doi.org/10.1093/scipol/scac081

Lepori, B., Jongbloed, B., & Hicks, D. (2023). Introduction to the Handbook of Public Funding of Research: Understanding vertical and horizontal complexities. In *Handbook of Public Funding of Research* (pp. 1–19). Edward Elgar Publishing.

https://www.elgaronline.com/edcollchap/book/9781800883086/book-part-9781800883086-5.xml
Matthias, L. (2020). Publisher OA Portfolios 2.0 (Version 2.0.) [Data set]. Zenodo. https://doi.org/10.5281/zenodo.3841568
Moed, H., Burger, W., Frankfort, J., & Van Raan, A. (1985). The application of bibliometric indicators: Important field-and time-dependent factors to be considered. *Scientometrics*, 8(3-4), 177-203.
Molas-Gallart, J. (2012). Research governance and the role of evaluation: A comparative study. *American journal of evaluation*, 33(4), 583-598.
Mongeon, P., & Paul-Hus, A. (2016). The journal coverage of Web of Science and Scopus: a comparative analysis. *Scientometrics*, 106, 213-228.
Morrison, H. (2019). OA APC longitudinal study dataset 2019 (Version 1.) [Data set]. Borealis. https://doi.org/10.5683/SP2/0DIPGE
Morrison, H. (2021a). 2011—2021 OA APCs (Version 1.) [Data set]. Scholars Portal Dataverse. https://doi.org/10.5683/SP2/84PNSG
Pinfield, S., Salter, J., & Bath, P. A. (2015). The "total cost of publication" in a hybrid open-access environment: Institutional approaches to funding journal article-processing charges in combination with subscriptions. *Journal of the Association for Information Science and Technology*, (In press). Retrieved from http://eprints.whiterose.ac.uk/81227/
Plan S. (2023). https://www.coalition-s.org/plan_s_principles/ . Retrieved 30/10/2023
Pollock, D., & Michael, A. (2019). Open access mythbusting: Testing two prevailing assumptions about the effects of open access adoption. *Learned Publishing*, 32(1), 7–12. https://doi.org/10.1002/leap.1209
Quaderi, N., Hardcastle, J., Petrou, C., & Szomszor, M. (2019). *The Plan S footprint: Implications for the scholarly publishing landscape*. Institute for Scientific Information Working Paper. Available at: https://beopen-project.eu/storage/files/ws190021-isi-report-2019-013.pdf
Reale, E., 2017. Analysis of National Public Research Funding (PREF) - Final Report, Publications Office of the European Union, Luxembourg, ISBN 978-92-79-73407-6, doi:10.2760/19140
Rodríguez, J. M., & Guns, R. (2022). Availability of Open Access journals by scientific fields, specialization and Open Access regulations in the YERUN universities. Transinformação, 34, e210064. https://doi.org/10.1590/2318-0889202234e210064
Ross-Hellauer, T. (2022). Open science, done wrong, will compound inequities. *Nature*, 603(7901), 363–363. https://doi.org/10.1038/d41586-022-00724-0
Ross-Hellauer, T., Reichmann, S., Cole, N. L., Fessl, A., Klebel, T., & Pontika, N. (2022). Dynamics of cumulative advantage and threats to equity in open science: a scoping review. *Royal Society Open Science*, 9(1), 211032.
Sastrón-Toledo, P; Alonso-Alvarez, P; Mañana-Rodriguez, J (2023). The cost of Open Access: public funding investment in APCs in Spain. *19th International Conference on Scientometrics & Informetrics, July 2-5, 2023 Bloomington, Indiana, USA.*
Solomon, D. J. (2014). A survey of authors publishing in four megajournals. *PeerJ*, 2, e365. https://doi.org/10.7717/peerj.365
Solomon, D. J., & Björk, B.-C. (2012a). A study of open access journals using article processing charges. *Journal of the American Society for Information Science and Technology*, 63(8), 1485–1495. https://doi.org/10.1002/asi.22673
Solomon, D. J., & Björk, B.-C. (2012b). Publication fees in open access publishing: Sources of funding and factors influencing choice of journal. *Journal of the American Society for Information Science and Technology*, 63(1), 98–107. https://doi.org/10.1002/asi.21660

Swan, A., & Brown, S. (2004). Authors and open access publishing. *Learned Publishing*, 17(3), 219–224. https://doi.org/10.1087/095315104323159649

Tananbaum, G. (2003). Of wolves and and boys: the scholarly communication crisis. *Learned Publishing*, 16(4), 285–289. https://doi.org/10.1087/095315103322422035.

The pandas development team. (2023). pandas-dev/pandas: Pandas (v2.1.3). Zenodo. https://doi.org/10.5281/zenodo.10107975

Whitley, R., 2000 [1984]. *The Intellectual and Social Organization of the Sciences*. Clarendon Press, Oxford.

Zacharewicz, T., Lepori, B., Reale, E., & Jonkers, K. (2019). Performance-based research funding in EU Member States—A comparative assessment. *Science and Public Policy*, 46(1), 105–115. https://doi.org/10.1093/scipol/scy041