

Opening the Black Box of Scholarly Communication Funding: A Public Data Infrastructure for Financial Flows in Academic Publishing

Stuart Lawson¹, Jonathan Gray², Michele Mauri³

[Abstract](#)

[1. Introduction](#)

[2. Towards a Systemic Picture of Financial Flows in Scholarly Communication](#)

[3. A Data Infrastructure for Journal Publication in the UK?](#)

[3.1 Institutional Income in the UK:](#)

[3.2 Subscriptions](#)

[3.3 Article Processing Charges \(APCs\)](#)

[4. Conclusions](#)

[Acknowledgements](#)

[Competing Interests](#)

[References](#)

Abstract

“Public access to publicly funded research” has been one of the rallying calls of the global open access movement. Governments and public institutions around the world have mandated that publications supported by public funding sources should be publicly accessible. Publishers are experimenting with new models to widen access. Yet financial flows underpinning scholarly publishing remain complex and opaque. In this paper we present work to trace and reassemble a picture of financial flows around the publication of journals in the UK in the midst of a national shift towards open access. We contend that the current lack of financial transparency around scholarly communication is an obstacle to evidence-based policy-making – leaving researchers, decision-makers and institutions in the dark about the systemic implications of new financial models. We conclude that obtaining a more joined up picture of financial flows is vital as a means for researchers, institutions and others to understand and shape changes to the sociotechnical systems that underpin scholarly communication.

Keywords: Scholarly Communication, Academic Publishing Finance, Open Access, Article Processing Charges (APCs), Higher Education Finance, Financial Transparency.

¹ Birkbeck College, University of London

² University of Amsterdam and Open Knowledge

³ Density Design, Politecnico di Milano

1. Introduction

Scholarly communication is a term generally used to refer to the formal process of publishing the results of research. Scholars and researchers communicate using many different channels but the term 'scholarly communication' is often predominantly used to refer to the "formal" written outputs (books, journal articles, conference papers, etc.) – though we are also currently seeing significant shifts in the "informal" channels through which research is mediated and circulated in society, from commercial platforms to informal sharing practices (Morris et al., 2013: 394).⁴ Responsibility for turning these research outputs into usable elements of the scholarly record has long been shared between various actors, with researchers providing the intellectual work and the validation system of peer review, publishers working on production, and librarians assuming responsibility for collecting, preserving, and facilitating access to works. This article will focus on journal articles rather than other research outputs because that is the area in which the fastest progress is being made and the area for the most relevant data is currently available.

The transition from printed materials to digital networked publishing has led to great changes in the scholarly communication process, not least of which is the ability to provide access to online copies of a work at near-zero marginal cost (Suber, 2012: 44). Combined with the global reach of the internet, the potential to communicate scholarship to a significantly broader audience than was possible under a print-dominated system has led to a flourishing of new ideas, tools, and organisations trying to make full use of the opportunities now available. However, we are currently operating in a mixed environment of print and digital, open and closed. All indications suggest that the trends from print to digital and from closed to open will continue, although the rate at which this is happening and the final outcome are difficult if not impossible to predict.

Open access is when research outputs are made available online free of cost to access and free of most restrictions on copying and reuse (Suber, 2012). In principle this enables anyone with the ability to access the internet to read and use these outputs. Even allowing for the inequalities inherent in citizens' ability to access online information due to social, economic, and language barriers, open access is clearly a great expansion of the potential readership for any given piece of academic work. The two main ways of providing open access to journal articles are through archiving work in repositories or by making work openly accessible at the point of publication through open access journals (known as 'gold' open access). Institutional and subject repositories (enabling 'green' open access) are also important parts of the research publishing system and several million articles are now available this way (Archambault et al., 2013). However, the economics of open access archiving are quite different, so this article will focus only on gold open access publishing - and in particular the subset which is funded by article processing charges or 'APCs'.

⁴ These are topics that we are studying further in other forthcoming papers.

Recent controversies around open access and scholarly communication have been underpinned by debates about money. Who should pay for what? Who should get paid for what? If research is made freely available on the web, how do publishers get paid? To what extent do publishers need to get paid (and what value do they add)? Is it fair that multinational corporate publishers have profit margins to rival oil companies or technology giants, based on the free labour of academics – access to whose work has to be purchased back by the institutions who supported them in producing it? If research is paid for by the public, shouldn't the public have access? If subscriptions won't pay for publishing, what will? Are new article processing charges for academics a new source of institutional inequality – or even a threat to academic freedom?

In this paper we argue for the importance of public systems for taking measure of the financial flows that underpin scholarly communication. We present some suggestions for making these flows visible. Taking the financing of journal articles from UK higher educational institutions as our point of departure, we argue that there is currently a dearth of public information that would be needed to enable evidence-based deliberation about the different ways of organising and paying for systems of scholarly communication. Based on our study of journal publication in the UK, we propose a provisional visual model as a starting point for analysing financial transparency around scholarly communication. We conclude that such a model could be useful not only for obtaining a better understanding of national systems, but also for facilitating trans-national comparisons and informing advocacy and policy work around scholarly communication systems which frequently transgress national borders.

Before proceeding into our analysis of financial flows around journal publication in the UK we will briefly pause to survey the broader context of this project. We are currently in the midst of numerous controversies and changes regarding the way that research circulates in society. Major research institutions say they can no longer afford the costs to provide their researchers with the material they need – giving rise to what has been described as the “serials crisis” (McGuigan, 2004; Panitch & Michalak, 2005; Young, 2009). Academics are boycotting publishers over what they consider to be excessive profiteering in the industry (Arnold & Cohn, 2012; Gowers, 2012; Marshall et al., 2015: 200-222; Neylon, 2012; The Economist, 2012). New digital technologies are changing the ways in which research can be circulated in society – from new open access publication models, to institutional repositories for researchers to self-archive their research, to new informal sharing practices (Eve, 2014; Gardner & Gardner, 2015; Suber, 2012).

In response to these and other developments, publishers, governments, funders, researchers and civil society groups are advocating and experimenting with new models for publishing research and for financing the costs of scholarly communication. Subscription-based publishing models are being complemented with other models in which researchers, institutions and funding councils pay for the costs of publication in different ways – such as by paying APCs.

What effect are these changes having on the way in which scholarly publishing is financed? Subscription journal publishing is a large global industry with the English-language Science, Technical, and Medical (STM) journals market estimated to be worth at least US\$10bn a year (Ware & Wabe, 2015). The costs associated with running open access journals are broadly similar to subscription journals (Wexler, 2015), although additional costs exist for those subscription journals which still produce print editions. However, the direct sources of *revenue* for publishers could be viewed as either very different – particularly under the APC model with individual researchers becoming involved in the payment process; or very similar – the funding for both subscription and open access articles (including APCs) usually originates with research institutions and their funders. The route by which this money makes its way into the hands of publishers does superficially look very different and this fact is causing a great deal of consternation among various stakeholders. The complex financial landscape we are currently faced with in scholarly communication will be outlined in detail below (see Figure 1).

In this paper we argue that it is currently difficult to analyse and evaluate the effects of new models for funding academic research due to the complex and opaque network of financial flows between public bodies, higher educational institutions, research councils, researchers and publishers. We propose a framework for mapping financial flows around scholarly communication, illustrated with reference to the financing of journal publications in the UK in the midst of a transition to 'gold' open access (i.e. the model whereby research outputs are made openly available at the point of publication). We contend that the current lack of financial transparency around scholarly communication is an obstacle to evidence-based policy – leaving researchers, decision-makers and institutions in the dark about the systemic implications of new models. We conclude that obtaining a more joined up picture of financial flows is vital as a means for researchers, institutions and others to understand and reshape the system intended to enable research to thrive.

2. Towards a Systemic Picture of Financial Flows in Scholarly Communication

Our enquiry into the finances of scholarly communication began with the question: What information is needed in order to obtain a bigger, more systemic picture of financial flows in scholarly communication? In particular we were interested in looking beyond the institutional level to understand how money flows from public institutions and funding sources to publishers. To what extent can one start to piece together such a picture using existing publicly available sources? And where are the gaps?

We propose that a systemic picture of financial flows around scholarly communication is essential for evidence-based discussion about not only about current arrangements (e.g. the overall effects of the APC model in the UK) – but also about how things might be organised

differently. This necessitates going beyond an institutional lens, and looking collectively at how much money is going into the system, where this money comes from, and how these financial flows might be adjusted to support different kinds of publishing models. It is our hope that such a systemic view might help to cut through an atmosphere of inevitability surrounding the current model (a world in which “there are no alternatives” as Margaret Thatcher put it) – and open up the imagination of researchers, policy-makers, librarians, open access advocates and others to new structural arrangements that could be supported through the collective investment of institutional funds.

The predominantly market-based organisation of academic publishing in the UK and many countries means that institutions look at costs and prices on a case by case basis, rather than have a bigger picture which could facilitate more effective collective coordination and decision-making. While some national bodies may possess a more comprehensive overview, this is not always complete and not always shared with researchers and institutions. In the medium term this situation could be addressed through a public data infrastructure which would enable the assembly and organisation of information about the finances around academic publishing. In another paper one of us has used the term “participatory data infrastructures” to describe sociotechnical systems for the production and circulation of information which are designed to be more attuned and responsive to the needs and concerns of their publics (Gray and Davies, 2015). In this context such a public data infrastructure might provide institutions, researchers, librarians, policy-makers, presses and others with information, indicators and evidence about the finances of academic publishing to enable informed interventions to shape the future of scholarly communication systems. We recognise, of course, that financial flows are only one element amongst a complex and contingent web of actors which comprise these systems (cf. Eve, 2014). Nevertheless, the prospect of the potential reinvestment and reallocation of public funds is a potentially a powerful collective mechanism for reshaping how scholarly communication systems are composed. Our work might thus be seen as an attempt to render visible resource flows in the service of more ambitious efforts in this direction.

To piece together a picture of these flows, we started by looking at what information we had from different sources in the UK (Figure 1b), and then abstracting this to a more general model which might help to inform and structure similar investigations in other countries (Figure 1a).

Figure 1a: Model of Financial Flows in Scholarly Publishing

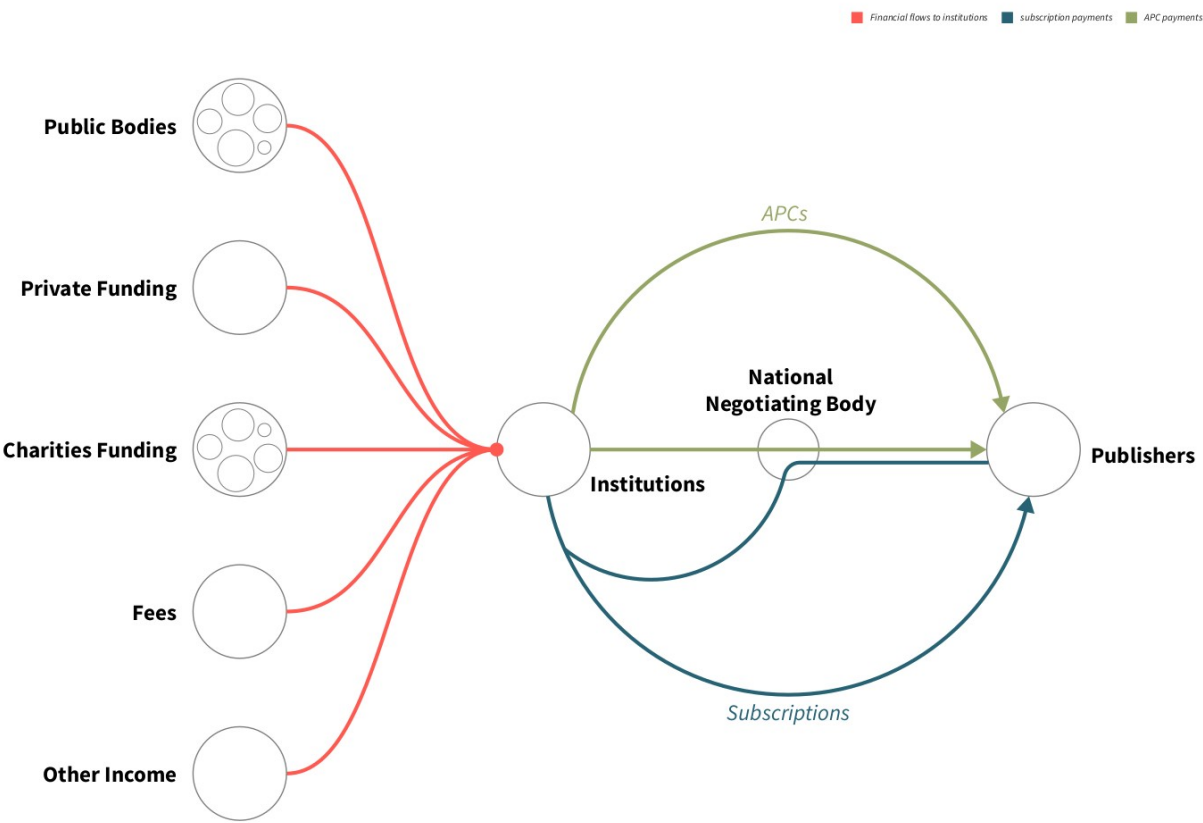
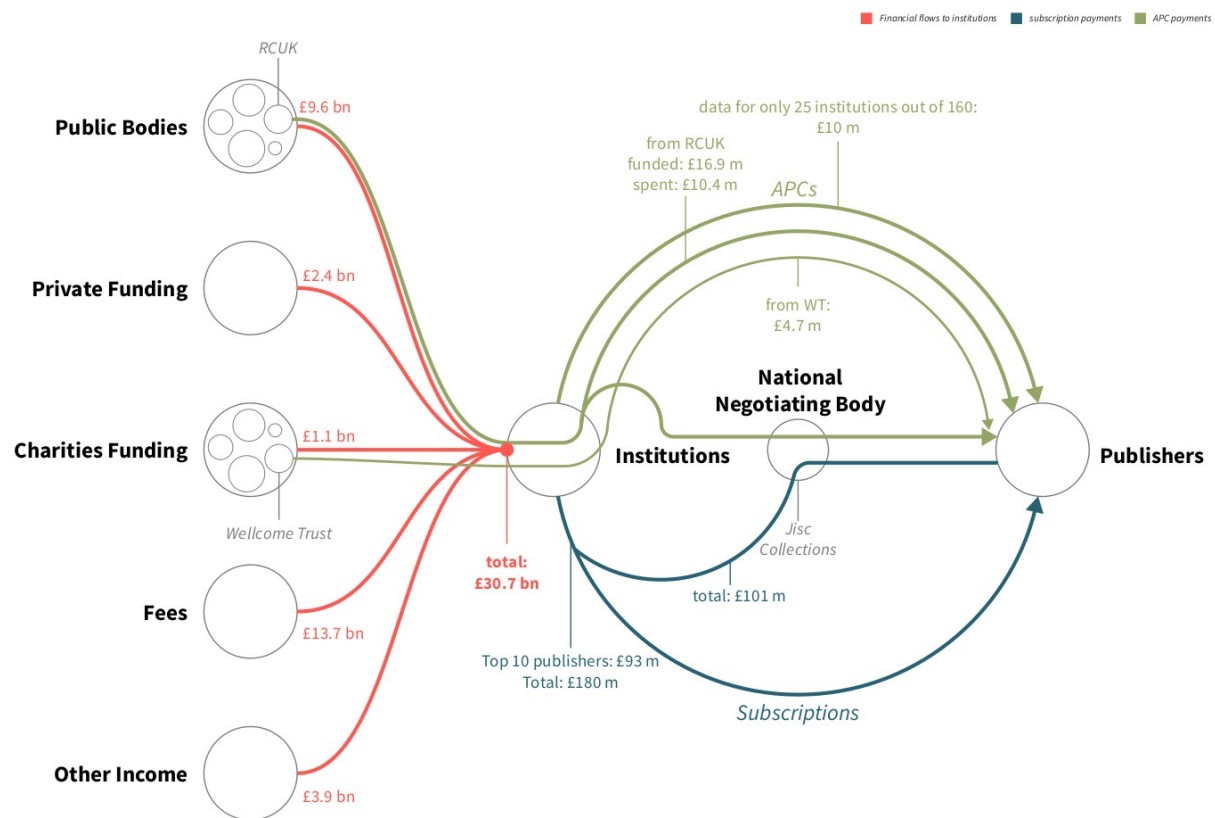


Figure 1b: Model of Financial Flows in Scholarly Publishing for the UK, 2014



The three main flows outlined in Figure 1 are various financial flows to institutions (orange), and then two flows from institutions to publishers: subscription payments (blue) and APC payments (green). The amounts shown in Figure 1b are for the year 2014. We also include another actor – national negotiating body – which is an organisation, often a library consortia, found in many (but not all) countries that acts as an intermediary negotiating on behalf of the library sector when purchasing access to journals (Gillies, 2014). Some elements are still missing from this picture. Subscription agents are organisations that handle some payments between libraries and publishers, particularly for small publishers, so that libraries do not have to deal with a large number of small publishers themselves. Higher education institutions also sometimes invest in publishers for example through direct shareholding or pensions schemes. The model also does not factor in VAT, which in the UK is charged on electronic publications (and APCs) but not print.

While our model omits some details, it aspires to render visible some of the main financial flows – as well to show where information is missing from public view. The following section examines the situation in the UK and is divided according to three flows outlined in the diagram: institutional income, institutional expenditure in the form of subscription payments, and institutional expenditure in the form of APC payments. As we shall see below, several crucial

financial flows – notably the full extent of either subscription or APC payments – are currently not publicly available in the UK. We hypothesise that this information asymmetry may provide a strategic advantage to big publishers in determining prices and profit margins, and a strategic disadvantage to public bodies, institutions and researchers advocating for publishing models which privilege public access over profit.

3. A Data Infrastructure for Journal Publication in the UK?

3.1 Institutional Income

The finances and practices of scholarly communication cannot be given due consideration without also understanding the broader contexts of higher education policies and the financing and organisation of higher educational institutions. In the UK, higher education is financed by a complex mix of public and private funding, which has repercussions for how we think about the flows of money in relation to scholarly communication. This section will describe the higher education funding situation at the time of writing in November 2015 with the caveat that further reforms will soon be underway, as outlined in a green paper from the Department of Business, Innovation and Skills (BIS, 2015).⁵

Public funding of UK higher education is provided via the BIS which funds both the higher education funding councils – led by the Higher Education Funding Council for England (HEFCE) – and the seven discipline-specific research councils (collectively known as RCUK). Since the reforms begun by the Conservative-led coalition government of 2010-15 – and extended by the current Conservative government – the system has been transitioning towards higher levels of private funding (McGettigan, 2013). Statistics collected annually by the Higher Education Statistics Agency (HESA) show how the proportion of higher education institution funding derived from public and private sources has been shifting as student tuition fee income has replaced teaching grants from the funding councils, a process that largely took place from 2012-15 as the government raised tuition fees for new undergraduate students from £3,225 to £9,000 starting from 2012 (Bolton, 2015) and withdrew the recurrent teaching grant for humanities subjects. In the academic year 2013/14 higher education institutions in the UK received £30.7bn of which £13.7bn (44.5%) was from tuition fees and £6.1bn (19.8%) was from funding body grants (HESA, 2015).

In Figure 1 tuition fees have been classed separately from either public or private funds because the precise legal status is unclear. Tuition fees for overseas students are 100% private, but it is more complex for home and EU students if they are funded by the state-backed loan scheme. It is currently estimated that around 45% of the state loans will not be repaid by students (McGettigan, 2015), so the total figure could perhaps be reasonably split 50:50 between public and private funds. Rather than use estimates in this way – the 45% figure has

⁵ If measures outlined in the Green Paper come into effect then further precarity will be introduced into our ability to collect the necessary data, with regards to both institutional income data (what if HESA is abolished?) and subscription data (what if Freedom of Information law no longer applies to universities?)

already been revised from previous estimates – for the purposes of the current model we have decided to retain tuition fees as a separate income stream. Higher education is a devolved matter in the UK so Scotland, Wales, and Northern Ireland all have different arrangements, particularly when it comes to charging tuition fees – research funding is still centrally determined by the UK government – but the devolved nations are included in the HESA statistics used to generate the amounts given in Figure 1b.

HEFCE provides various streams of funding to institutions, with the largest elements being teaching grants for STEM subjects – the recurrent teaching grant for humanities and social science subjects (or any disciplines not designate as 'high cost subjects') were phased out beginning 2012 (HEFCE, 2012) – and research grants in the form of quality-related (QR) research funding, which is allocated according to institutions' performance in the Research Excellence Framework (REF). QR funding totalled £1.6bn in 2015/16 (HEFCE, 2015). An additional £2.67bn also originated from BIS in the form of Research Council grants (RCUK, 2015).

Institutions also receive income from a variety of other sources, such as events, intellectual property exploitation, fundraising etc. Two significant sources of research funding for UK higher education institutions are medical charities and the European Union. Among medical charities the Wellcome Trust plays a large role; it spent £674m in research grants and other charitable activities in 2014 (Wellcome Trust, 2015), with a majority of this going to UK-based researcher projects (personal communication – Wellcome Trust, 2015). The EU funds a programme of research activities via the multi-year Framework Programmes organised by the European Commission; we are currently in Framework Programme 8, known as Horizon 2020, covering expenditure for the years 2014-2020.

An interesting question is whether it is possible to trace the university funding and different income streams in what is paid out to publishers. According to the *Further and Higher Education Act 1992* and HEFCE's *Memorandum of assurance and accountability between HEFCE and institutions* (2014), institutions must use grants received from the UK's funding bodies for the prescribed purposes (i.e. teaching and research) but the main teaching grant and QR research grant go into an institution's general funds so cannot be precisely traced from income to expenditure. In other words, money used by an institution to pay for journal subscriptions may originate from a combination of multiple sources e.g. tuition fees, HEFCE grants, endowments etc. This means that while it is possible to calculate the proportion of institutional income from public funds, it is also not possible to state the proportion of money expended on journal subscriptions to come from public funds.

The situation is somewhat different for APCs. In 2011 David Willetts, the Minister for Science and Universities at the time, commissioned a working group led by Janet Finch to look into the possibility of transitioning towards open access. The resulting report — commonly known as the Finch Report (Finch Group, 2012) — made various policy recommendations designed to encourage greater uptake of APC-funded gold open access. RCUK acted on these

recommendations by introducing an open access policy requiring all research that they fund to be made open access (RCUK, 2013). Full compliance was not expected immediately; RCUK has provided block grants to institutions in order to pay for APCs covering five years from 2013/14, with the expectation that by Year Five of the policy (2017/18) 75% of RCUK-funded articles must be made available through immediate gold open access (RCUK, 2013a). The funds provided to institutions to support this policy started with £16.9m in the academic year 2013/14 (RCUK, 2014) and rise slightly for each year of the policy. Institutions must report back to RCUK with data on APC expenditure in order to monitor compliance (see *Article processing charges (APCs)* section below) so there is a strong accountability mechanism in place to trace the flow of this money.

Thanks to figures available from HESA – which unfortunately are not openly available and require payment in order to access – institutional budgets and their income sources can be scrutinised in full, making this the most transparent element of the system as visually represented in Figure 1. It is worth noting that there are considerable structural difference between national higher education systems in terms of the balance of public and private financing, and some are largely comprised of private institutions that are subject to different transparency and accountability standards. Movement of money *within* UK institutions is more opaque because while they all publish annual financial reports, for the most part we only know about the income and expenditure streams rather than the internal distribution of funds. This will be evident from the following examination of two expenditure flows from institutions – subscriptions and APCs.

3.2 Subscriptions

Around 35,000 peer-reviewed academic journals are currently being published worldwide (Ware & Wabe, 2015: 6), of which over 10,000 are open access (DOAJ, 2015). The market is estimated to be worth around US\$10bn (Ware & Wabe, 2015) with academic and research libraries being publishers' biggest customers. Many consider it to be a dysfunctional market (Cockerill, 2006; Shieber, 2009) which is of particular concern when much of this money is from public funds. Journal subscriptions are usually paid for out of institutional library budgets; other sources such as personal subscriptions and pay-per-view make up a very small proportion of publisher revenues. Library expenditure comprises approximately 2.5% of UK university budgets⁶. As with their parent organisations the largest item of expenditure for libraries is staff, followed by resources expenditure, of which electronic journal subscriptions make up the largest element (Research Information Network, 2010).

It has historically been difficult to know the precise amounts which are paid by institutions for journal subscriptions. If we consider that full transparency of the flows of money within the system would require knowing how much each institution pays to each publisher for each journal, there are multiple factors which have made it difficult to gain this knowledge. The first is that libraries have not historically published accounts of their financial expenditure at such a

⁶ This figure is based on data obtained from SCONUL (Society of College, National and University Libraries) which is not open.

fine-grained level of detail; even if it is possible to know a library's total serials budget for a given year (such as through the annual collection of library statistics by SCONUL in the UK), this figure is not broken down into individual line items to specific publishers. The second factor is the bundling of journal subscriptions into packages known as 'big deals', whereby a library will purchase access⁷ to a collection of titles by a publisher for a single sum rather than paying individual subscription rates.

Bundling has led to obscuring the costs in two distinct ways. The first is that it means the individual list price of a journal subscription – which can usually be easily found on publisher websites – is often not the price that is actually paid for access to a title; in fact it may not be possible to disaggregate the big deal price in order to know how much was paid for a particular individual title. In the UK the majority of big deals are negotiated and administered by the library consortium Jisc Collections – part of Jisc, the technology infrastructure body for higher and further education – with 56% of institutions' serials expenditure going on Jisc Collections deals.⁸

Some big deal prices still take into account the amount an institution was paying when it first signed up, which could have been more than 15 years ago. This 'historical print spend' calculation leads to some institutions paying far more than others – even institutions of a similar size with a similar library budget – to access the same content. On the whole, wealthier institutions pay more, but this does not always hold true.

The second obscuring factor is the presence of confidentiality and non-disclosure clauses in contracts. Some publishers include such clauses in the licenses that libraries sign in when they purchase access to journal content. The extent to which this practice goes on is unclear but it appears to be widespread globally. In the UK, in order to introduce an element of clarity to the situation, a number of researchers have used Freedom of Information (FOI) requests to make journal expenditure information public. Lawson and Meghreblan (2015) describe the methods used to get data for payments to publishers by UK higher education institutions in a data article; *Table 1* shows a summary of this data.

	2010	2011	2012	2013	2014
Elsevier	£34,177,020	£36,781,827	£39,079,332	£39,476,813	£39,812,145
Wiley	£13,460,226	£14,662,250	£15,616,311	£16,369,917	£16,875,190
Springer	£7,311,046	£7,309,094	£7,906,177	£7,940,116	£8,542,997
Taylor & Francis	£8,319,095	£9,140,572	£9,710,528	£10,084,350	£10,828,334
Sage	£4,495,313	£5,085,196	£5,608,296	£5,869,791	£5,990,818

⁷ Note the use of the term 'access'. When purchasing online journal content, a customer is not necessarily buying ownership but rather the rights of access.

⁸ This figure is based on information obtained by one author (Lawson) who previously worked at Jisc Collections.

Oxford University Press	£1,996,163	£2,163,242	£2,395,136	£2,669,757	£2,925,607
Cambridge University Press	£1,447,978	£1,462,214	£1,690,078	£1,832,177	£1,885,485
Nature Publishing Group	£2,998,040	£3,593,308	£4,066,962	£4,273,822	£4,430,900
Royal Society of Chemistry	£806,129	£867,752	£1,062,237	£1,062,948	£1,101,860
Institute of Physics Publishing	£1,091,517	£1,119,070	£1,197,958	£1,279,691	£1,373,533
Total for these 10 publishers	£76,102,528	£82,184,527	£88,333,015	£90,859,384	£93,766,870

Table 1: Subscription expenditure of UK higher education institutions with ten publishers, 2010-14

(Source: <http://dx.doi.org/10.6084/m9.figshare.1186832>)

The ten publishers in *Table 1* received £94m in subscription revenues from UK academic libraries in 2014. Since the total UK serials expenditure was around £180m in 2013/14 (according to SCONUL data⁹), there is still £86m that remains unaccounted for. While it would be technically possible to obtain all of this information through further FOI requests the burden placed on academic libraries by doing this would be great, so a more sustainable long-term solution would be to create an online service which libraries can voluntarily contribute their data to each year.

A similar approach to finding out costs has been attempted in a few other nations as well such as Switzerland (Gutknecht, 2014) and New Zealand (Wilson, 2014) albeit with limited success. In many countries the majority of payments to publishers are made by higher education institutions which are partly or fully funded by public money, so FOI requests can be used to make public a significant amount of the payments within the global scholarly communications market.

Open access advocacy has often been motivated in part by the large amounts of money which we see flowing from public funding to a handful of large corporations (see Larivière, Haustein & Mongeon (2015) for more on the oligopolic nature of the current academic journals market). When combined with the questionable level of added value created by publishers – which may be non-negligible but most of the labour is undertaken by academics themselves – then 35-40% profit margins and billion-pound yearly profits are considered extremely galling. This can be critiqued as signalling the extent to which universities are functioning as sources of

⁹ SCONUL have changed the terms of their data collection and as of academic year 2014/15 the serials expenditure category is no longer obligatory (SCONUL, personal communication), so an equivalent total figure will not be available for subsequent years.

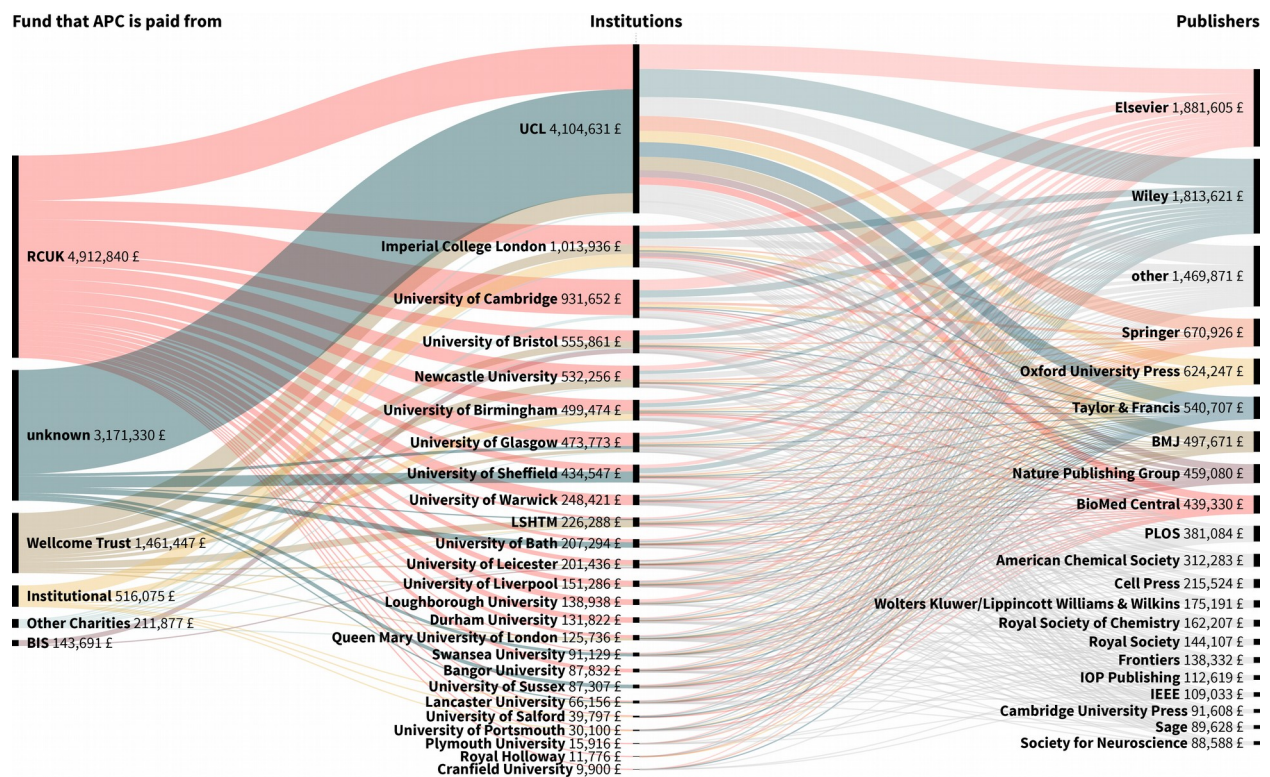
profit to be extracted for maximum value rather than as sites of critical resistance to forms of domination. On the other hand, it is worth noting the argument that making expenditure data open could simply act in the service of making markets more “efficient” – in direct accordance with neoliberal rhetoric – rather than as a means to pursue public interest structural reforms to scholarly communication.

3.3 Article Processing Charges (APCs)

Making work openly accessible at the point of publication through open access journals can be funded in various different ways. Many journals, particularly in niche subjects in the humanities and social sciences, are run on low- or zero budgets and volunteer academic labour. Others are funded by institutional subsidies from a university, scholarly society, or research funder. Consortial funding models are also being explored by publishers such as Open Library of Humanities and Knowledge Unlatched. One method of funding journals that is gaining prominence is to charge a fee known as an article processing charge (APC) which is to be paid by the author's institution, funder, or sometimes out of their own pocket. According to the Directory of Open Access Journals (DOAJ) around a third of open access journals charge an APC (DOAJ, 2015a). This does not necessarily mean that only one third of open access articles are published after an APC is paid because these journals include many of the more high-volume open access journals and publishers, such as PLOS, and the majority of closed-access journals from subscription publishers are now 'hybrid' journals which include the option of paying an APC to make an article open access within an otherwise closed journal.

The amount of money being paid to publishers for APCs has been growing strongly for the past few years and this has largely been made possible by research funders making funds available specifically for this purpose. In the UK, as mentioned above, RCUK distributed block grants of varying amounts to 107 research institutions (RCUK, 2014). These funds are usually managed by either research support staff in the library or research administrators. In order to make sure that their money is being spent appropriately, each year institutions must report back to RCUK on the level of expenditure from these block grants in order to monitor compliance with the open access policy. Jisc has worked with RCUK to create a template for institutions to report their block grant APC expenditure in a standardised way and release it openly (Jisc Collections, 2015). This high level of scrutiny allows us an unprecedented amount of access to the inner workings of the APC market. Figure 2 uses this data to visualise the financial flows from funders to publishers via institutions.

Figure 2: APC payments by 25 UK institutions, 2014



In this dataset ‘Unknown’ is the second largest category of expenditure. The bulk of APCs of unknown origin can be attributed to University College London (UCL). The fact that much APC funding originates from RCUK, who allocate amounts according to past research performance, and that UCL – one of the ‘elite’ Russell Group universities – spent significantly more on APCs in 2014 than any other institution, raises multiple questions about the structure of the APC market and its relation to institutional reputation. Might a shift towards funding research publication via APCs reinforce existing inequalities between different institutions? How will this affect the relative ability of researchers from other nations to publish work in the same venues and through the same processes as UK researchers? In 2014 UCL were the second largest actor in this figure after RCUK – what happens to the APC market if they change their strategy and cut APC funding?

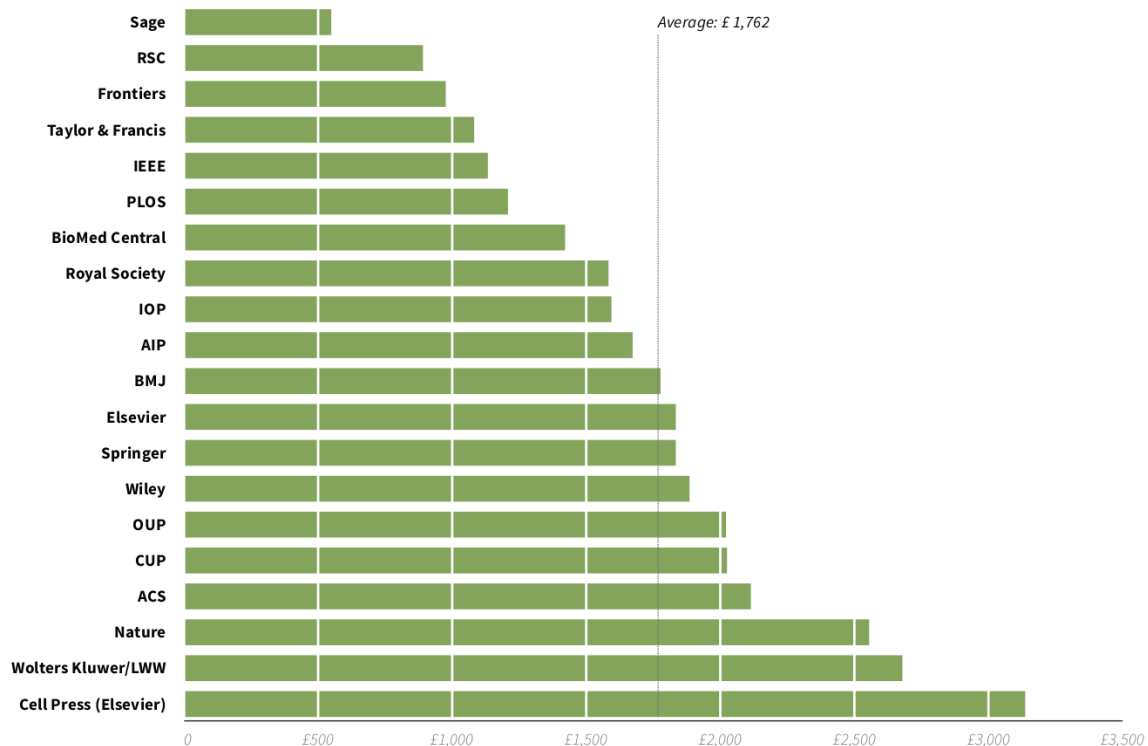
The fact that financial transparency at the micropayment level (i.e. APCs) is becoming an expected norm is a striking contrast from the historical situation in the subscriptions market, in which the limited level of transparency we currently have has only been achieved through using FOI law. Bringing the same scrutiny to bear on subscription costs by transferring the transparency principle to that area could hopefully lead to richer public discussion of the value of those subscriptions and ways of transitioning away from the model.

We have learned from this monitoring that the majority of APC payments arising from funder grants have been to hybrid journals (Jisc, 2014; Wellcome Trust, 2015a). This is perhaps unsurprising for two reasons. Firstly, because the majority of 'high impact' journals are subscription journals. And secondly, because few funders have set a cap on the maximum amount that can be spent on an individual APC – and APCs in hybrid journals are higher than those in full open access journals (Björk & Solomon, 2014; Pinfield, Salter & Bath, 2015). The available data is mostly on APCs paid by European funding agencies – details of over 10,000 APCs paid by around 50 UK higher education institutions are now openly available online (if a few sources are combined e.g. Lawson 2015b and Lawson 2015c) and a growing number from German (Apel et al., 2015; Sikora & Geschuhn, 2015) and Austrian (Reckling & Rieck, 2015) institutions as well – but we know significantly less about APCs paid by institutions (or even authors) themselves and this is a limitation of any analysis based on currently available data.

An example from the UK can illustrate this point. Jisc Collections has designed a standardised template for UK higher education institutions to use for recording APC data and now collects this data for analysis and makes it openly available in the online data archive figshare (see Lawson 2015a for more details). Data is available from 26 institutions for 2014, and 15 institutions covering the period 2013-15 (at time of writing figures for 2015 only cover the six months January – June)¹⁰. Figure 3 shows the average APC price paid by UK universities to various different publishers during the first half of 2015. These prices vary but the average figure is around £1,700 which is similar to 'average APC' estimates from other sources (Björk & Solomon, 2014; Pinfield, Salter and Bath, 2015). This is the best available data we have on actual APCs that are paid, as opposed to list prices which are what Björk and Solomon based their estimate on. Therefore it would seem reasonable to use this figure as a benchmark. However, that would be misleading for the following reasons.

¹⁰. Full year data for 2015 will be made available in February 2016.

Figure 3: Average APC price paid to publishers by UK institutions, 2015



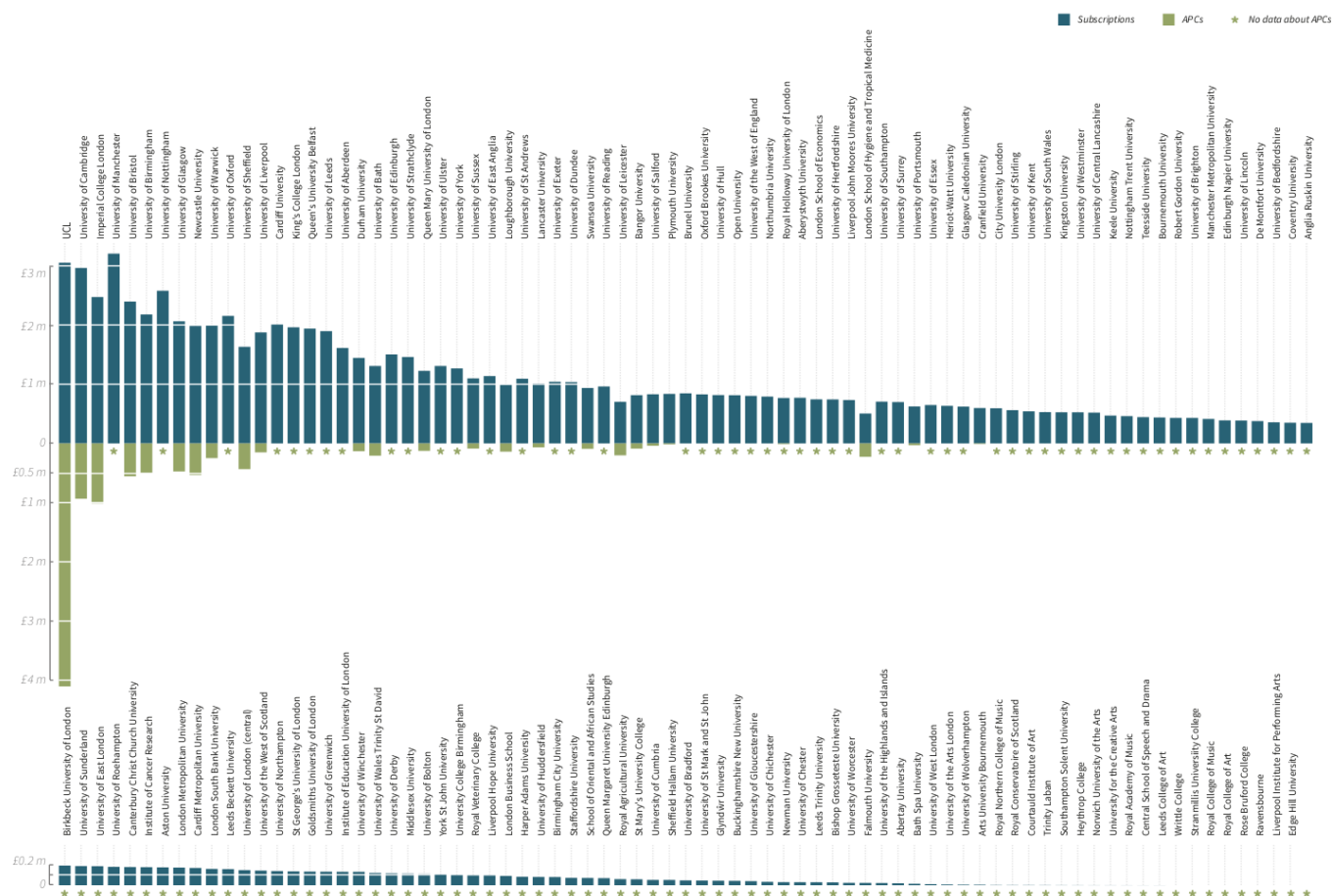
(source: http://figshare.com/articles/Average_APC_price_2015/1509861)

As noted above, the majority of gold open access journals do not charge APCs and while it is not yet possible to know precisely the percentage of open access articles for which an APC was required, it would be wrong to assume that most do. It would also be disingenuous to think of the APC market as a single market converging on one price point. An analysis of APC price information collected by DOAJ shows the wide variation in list prices of APCs [also Björk and Solomon; Pinfield, Salter and Bath]. There is some competition among new entrants such as PeerJ, SAGE Open, and Ubiquity Press, which all have drastically lower costs than the stated average. We may also be seeing the emergence of a competitive market for APCs at the top end among the 'highest impact' journals (as measured by the much abused Impact Factor metric). Recent data covering APCs paid during the first half of 2015 in the UK shows that two publishers with strong 'prestige' brands, Nature Publishing Group and Cell Press, are showing above-average growth despite their above-average APC prices (analysis based on comparing data from Lawson 2015d and Lawson 2015e). The majority of known APCs are paid to hybrid journals and by extension we can see that established commercial publishers are dominating the APC market in a similar way to the subscription market (see Figure 2; also Pinfield, Salter and Bath, 2015).

A great deal about the APC market is still unknown. RCUK and Wellcome Trust funding is predominantly for STEM research so we have less data regarding APC expenditure in humanities and social sciences disciplines. In the UK, the vast majority of humanities research is unfunded – or at least, not *directly* funded, since HEFCE QR funding will be indirectly responsible for paying some researchers. A further gap in current knowledge is APCs which are paid by UK academic researchers from funds which are not managed centrally. Preliminary research undertaken by Stephen Pinfield regarding APC payments at the University of Nottingham indicates that a significant percentage of APCs go unrecorded because they are paid for out of individual department or project funds (personal communication). Figure 4 shows the combined expenditure on subscriptions and APCs by a number of institutions in 2014 and clearly shows the gaps in current knowledge. For many institutions we just do not know how much they spent on APCs. For example, University of Oxford is displaying zero APC expenditure because the data is not available, but we know they spent £800,000 on APCs in the first six months of 2015 alone (Lawson & Evans, 2015). And even the APC expenditure which *is* reported is only that which is centrally paid for by the institution so is incomplete. The amounts spent on subscriptions are also incomplete because they only cover the 10 largest publishers.

Returning for a moment to the broader theme of this article – tracing financial flows and placing them within a larger systemic context – we can see a mixed picture with regards to our current ability to understand how all the pieces fit together. While the available data reveals a lot of detail – regarding which institutions are receiving money for APCs thanks to RCUK block grant allocations, and which publishers they are paying thanks to openly published APC expenditure data – a great deal remains unknown so a systemic picture of the financial flows around APCs is currently somewhat fuzzy. The incompleteness we see here reinforces our proposal that a more robust open data infrastructure is needed to enable a fully-informed discussion around possible ways of organising the funding of research publication.

Figure 4: Expenditure on subscriptions and APCs in the UK, 2014



4. Conclusions

In this paper we argue for the importance of a joined-up, systemic, publicly accessible picture of financial flows around academic publishing to inform evidence-based deliberation, policy and action to shape scholarly communication systems. We have proposed a provisional visual model for assessing the availability of information about these financial flows, taking journal publication in the UK as a case study. Our analysis of three broad types of flows in this model – institutional income, subscription payments and APCs – highlights that there are still significant obstacles to obtaining the information that is needed to piece together a bigger picture.

There is systemic opacity both within institutions as well as regarding the “black box” of finances around scholarly communication in the UK as a whole. Just as we do not yet possess an accurate overview of how much the UK’s total APC expenditure is, institutions often lack aggregate figures of their total APC payments towards a given publisher across different departments.

Why does this matter? Or – to frame the question a different way – who might stand to benefit from this opacity, and who might lose out? This question cannot be reflexively posed without at least minimally considering broader shifts towards transparency, reporting, quantification and “audit rituals” within public institutions – as well as shifts in management and governance within higher education (cf. Power, 1999; McGettigan, 2013; Davies, 2014). A higher degree of transparency around financial flows in scholarly communication might have consequences for institutions of higher learning whose members often complain of the crippling explosion of administrative rituals in addition to excessive teaching loads and diminishing research time. The Conservative government in the UK has very consciously adopted financial transparency as a key part of their strategy to pursue austerity in public finances – to increase pressure on public institutions to cut costs and increase efficiency gains (see Worthy, 2013; Gray, 2014). Any calls for greater transparency must be balanced against risks and unexpected consequences for researchers, institutions and other actors in the scholarly communication system.

Yet at the same time, the lack of a bigger picture about the effects of systemic changes in the financing of academic publishing also may inhibit researchers and institutions' ability to effectively advocate and take action in pursuit of their collective interests – and the interests of other actors who might stand to benefit from increased access to research. This might include collective negotiation over subscription prices and publication charges through national bodies such as Jisc. It might include other forms of collective intervention to address unfair practices by large publishers – such as the recent case of a country-wide boycott of Elsevier in the Netherlands (see Kingsley & Harnad, 2015; Wijkhuijs, 2015). A sharper empirical picture of the collective resources that institutions might also inspire greater experimentation with other financial models (such as the “consortial” models of Knowledge Unlatched or the Open Library of Humanities) or the pursuit of more structurally ambitious changes to public policy. We propose that further research in this area is needed.

While this article focuses on the UK, we would be very interested to see or collaborate around further work to obtain a systemic picture of financial flows around scholarly communication in other countries. We provide a provisional model abstracted from our analysis of UK journal publication finances (Figure 1a) that may serve as a starting point for comparison. We recognise, of course, that there is a high degree of variance between countries in how scholarly communication is funded, in the composition of the higher education sector and in the organisation of the public sector and public sector funding more generally. This situation is further complicated by several decades of “public sector reform” initiatives in the service of efficiency gains and marketisation. As transparency researcher David Heald comments, “the positive [welfare] state that has been displaced was much easier to comprehend, map and record than the successor regulatory state” (Heald, 2012: 41). However, despite these obstacles we believe that greater transnational comparability could be advantageous for institutions, researchers, policy-makers and collective negotiation bodies interested in advocating for a fairer system.

The current lack of publicly available information concerning financial flows around scholarly communication systems is an obstacle to evidence-based policy-making – leaving researchers, decision-makers and institutions in the dark about the implications of current models and the resources available for experimenting with new ones. Through our work in this paper we hope to have made a modest contribution towards a public data infrastructure to render these financial flows visible and accessible so that researchers, institutions and others are able to understand and shape changes to the sociotechnical systems that underpin scholarly communication.

Acknowledgements

We would like to thank Open Knowledge, the Digital Methods Initiative (University of Amsterdam), Centre for Technology and Publishing (Birkbeck, University of London) and Density Design (Politecnico di Milano) for hosting the research workshop which led to this paper. The workshop participants were Richard Rogers, Liliana Bounegru, Natalia Sanchez, Jonathan Gray, Erik Borra, Martin Eve, Stuart Lawson, Michele Mauri, and Matteo Azzi. Parts of this research were supported through a grant from the Open Society Foundations (OSF) as well as through the EU FP7 co-financed PASTEUR4OA (Open Access Policy Alignment Strategies for European Union Research) project.

Competing Interests

SL was previously employed by Jisc Collections, a library consortium involved in negotiations with publishers, which is mentioned in the article. SL is undertaking a PhD which is part-funded by Jisc Collections and supervised by the co-founders of Open Library of Humanities, Martin Eve and Caroline Edwards. All other authors have no competing interests.

References

Apel, J, Becker, H-G, Bertelmann, R, Beucke, D, Blume, P, Blumtritt, U, Busjahn, D, Deinzer, G, Dorner, A, Engelhardt, C, Hell, D, Herb, U, Feldsien-Sudhaus, I, Franke, F, Frick, C, Geißelmann, A, Geschuhn, K K, Kuehle, G, Jaeger, D, Kroiss, S, Lucht-Roussel, K, Lützenkirchen, F, Oberländer, A, Peil, V, Pieper, D, Pohlmann, T, Schlachter, M, Ruckelshausen, F, Schlegel, B, Sikora, A, Tullney, M, Vieler, A, Witt, S and Jahn, N (2015). *Datasets on fee-based Open Access publishing across German Institutions*. Bielefeld University. DOI: <http://dx.doi.org/10.4119/UNIBI/UB.2014.18>

Archambault, E, Amyot, D, Deschamps, P, Nicol, A, Rebout, L, and Roberge, G (2013). *Proportion of open access peer-reviewed papers at the European and world levels—2004-2011*. Brussels: Science-metrix; European Commission. Available at http://www.science-metrix.com/pdf/SM_EC_OA_Availability_2004-2011.pdf [Last accessed 28 October 2015].

Arnold, D N and Cohn, H (2012). Mathematicians take a stand. *Notices of the American Mathematical Society*, 59(6): 828-833. DOI: <http://doi.org/10.1090/noti857>

Björk, B-C and Solomon, D J (2014). *Developing an effective market for open access article processing charges*. London: Wellcome Trust. Available at https://www.fwf.ac.at/fileadmin/files/Dokumente/Downloads/Dev_Effective_Market_OA_Article_Processing_Charges.pdf [Last accessed 28 October 2015].

Cockerill, M (2006). The economics of open access publishing. *Information Services and Use*, 26(2): 151–157. Available at <http://content.iospress.com/articles/information-services-and-use/isu499> [Last accessed 28 October 2015].

Davies, W (2014). *The Limits of Neoliberalism: Authority, Sovereignty and the Logic of Competition*. London: Sage.

Department for Education (1992). *Further and Higher Education Act*. Available at <http://www.legislation.gov.uk/ukpga/1992/13/contents> [Last accessed 28 October 2015].

Department of Business, Innovation and Skills (2015). *Fulfilling our potential: Teaching Excellence, Social Mobility and Student Choice*. Available at https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/474266/BIS-15-623-fulfilling-our-potential-teaching-excellence-social-mobility-and-student-choice-accessible.pdf [Last accessed 6 November 2015].

Directory of Open Access Journals (2015). *About DOAJ*. Available at <https://doaj.org/about> [Last accessed 7 November 2015].

Directory of Open Access Journals (2015a). *Historical APC data from before the April upgrade*. Available at <https://doajournals.wordpress.com/2015/05/11/historical-apc-data-from-before-the-april-upgrade/> [Last accessed 7 November 2015].

Eve, M P (2014). *Open Access and the Humanities: Contexts, Controversies and the Future*. Cambridge: Cambridge University Press. DOI: <http://dx.doi.org/10.1017/CBO9781316161012>

Finch Group (2012). *Accessibility, sustainability, excellence: how to expand access to research publications. Report of the Working Group on Expanding Access to Published Research Findings*. Available at <http://www.researchinfonet.org/wp-content/uploads/2012/06/Finch-Group-report-FINAL-VERSION.pdf> [Last accessed 2 November 2015].

Gardner, C C and Gardner, G J (2015). Bypassing Interlibrary Loan Via Twitter: An Exploration of #icanhazpdf Requests. In: *Proceedings of ACRL 2015*, Portland, Oregon. Available at <http://eprints.rclis.org/24847/> [Last accessed 9 November 2015].

Gillies, S (2014). Negotiating author fees for hybrid OA journals: Library consortial leadership. *Collection Management*, 39(4): 231–252. DOI: <http://dx.doi.org/10.1017/CBO9781316161012>

Gowers, T (2012). Elsevier — my part in its downfall. Available at <https://gowers.wordpress.com/2012/01/21/elsevier-my-part-in-its-downfall/> [Last accessed 9 November 2015].

Gray, J (2014). Towards a Genealogy of Open Data. Available at SSRN: <http://dx.doi.org/10.2139/ssrn.2605828>

Gray, J and **Davies, T** (2015). Fighting Phantom Firms in the UK: From Opening Up Datasets to Reshaping Data Infrastructures? Available at SSRN: <http://dx.doi.org/10.2139/ssrn.2610937>

Gutknecht, C (2014). *Intransparenz bei den Bibliotheksausgaben von Schweizer Hochschulen*. wisspub.net. Available at <http://wisspub.net/2014/10/13/intransparenz-bei-den-bibliotheksausgaben-von-schweizer-hochschulen/> [Last accessed 7 November 2015].

Heald, D (2012). Why is transparency about public expenditure so elusive? *International Review of Administrative Sciences*, 78(1): 30–49. DOI: <http://doi.org/10.1177/0020852311429931>

HEFCE (2012). *Recurrent grants and student number controls for 2012-13*. Available at http://www.hefce.ac.uk/media/hefce/content/pubs/2012/201208/12_08_1123.pdf [Last accessed 2 November 2015].

HEFCE (2014). *Memorandum of assurance and accountability between HEFCE and institutions*. Available at http://www.hefce.ac.uk/media/hefce/content/pubs/2014/201412/HEFCE2014_12.pdf [Last accessed 28 October 2015].

HEFCE (2015). *Recurrent grants for 2015-16*. Available at http://www.hefce.ac.uk/media/HEFCE_2014/Content/Pubs/2015/201505/HEFCE2015_05.pdf [Last accessed 28 October 2015].

HESA (2015). *Income of UK HE providers*. Available at https://hesa.ac.uk/index.php?option=com_content&view=article&id=1900&Itemid=634 [Last accessed 28 October 2015].

Jisc (2014) *Jisc evidence to the independent review of the implementation of RCUK Policy on Open Access*. Available at <http://scholarlycommunications.jiscinvolve.org/wp/files/2014/09/Jisc-evidence-to-RC-OA-policy-review-20140910.docx> [Last accessed 10 November 2015].

Jisc Collections (2015). *APC data collection*. Available at <https://www.jisc-collections.ac.uk/Jisc-Monitor/APC-data-collection/> [Last accessed 4 November 2015].

Kingsley, D and **Harnad, S** (2015). Dutch universities plan Elsevier boycott — will this be a game changer or will publisher profits remain unaffected? *LSE Impact blog*. Available at <http://blogs.lse.ac.uk/impactofsocialsciences/2015/07/08/dutch-universities-boycott-elsevier/> [Last accessed 28 October 2015].

Larivière, V, Haustein, S and Mongeon, P (2015). The Oligopoly of Academic Publishers in the Digital Era. *PLOS ONE* 10(6): e0127502. DOI: <http://dx.doi.org/10.1371/journal.pone.0127502>

Lawson, S (2015). 'Total cost of ownership' of scholarly communication: managing subscription and APC payments together. *Learned Publishing*, 28(1). DOI: <http://dx.doi.org/10.1087/20150103>

Lawson, S (2015a). Article processing charges paid by 25 UK universities in 2014. *Journal of Open Humanities Data*, 1. DOI: <http://doi.org/10.5334/johd.2>

Lawson, S (2015b). APC data for 25 UK higher education institutions - 2014. *figshare*. DOI: <http://dx.doi.org/10.6084/m9.figshare.1305596>

Lawson, S (2015c). APC data for 27 UK higher education institutions - 2015. *figshare*. DOI: <http://dx.doi.org/10.6084/m9.figshare.1507481>

Lawson, S (2015d). Aggregated APC data 2014. *figshare*. DOI: <http://dx.doi.org/10.6084/m9.figshare.1311672>

Lawson, S (2015e). 2015 (Jan-June) UK APC data combined. *figshare*. DOI: <http://dx.doi.org/10.6084/m9.figshare.1509860>

Lawson, S and **Evans, A** (2015). University of Oxford APCs 2015 (Jan-June). *figshare*. DOI: <http://dx.doi.org/10.6084/m9.figshare.1506829>

Lawson, S and **Meghreblian, B** (2014). Journal subscription expenditure of UK higher education institutions. *F1000Research*, 3:274. DOI: <http://dx.doi.org/10.12688/f1000research.5706.3>

Marshall, J P, Goodman, J, Zowghi, D, and da Rimini, F (2015). *Disorder and the Disinformation Society: The Social Dynamics of Information, Networks and Software*. London: Routledge.

McGettigan, A (2013). *The Great University Gamble: money, markets, and the future of higher education*. London: Pluto Press.

McGettigan, A (2015). *The accounting and budgeting of student loans*. HEPI report 75. Available at <http://www.hepi.ac.uk/wp-content/uploads/2015/05/Accounting-and-Budgeting-FINAL.pdf> [Last accessed 7 November 2015].

McGuigan, G S (2004). Publishing Perils in Academe. *Journal of Business & Finance Librarianship*, 10(1): 13–26. DOI: http://doi.org/10.1300/J109v10n01_03

Morris, S, Barnas, E, LaFrenier, D and Reich, M (2013) *The Handbook of Journal Publishing*. Cambridge: Cambridge University Press.

Neylon, T (2012). The Cost of Knowledge. Available at <http://thecostofknowledge.com/> [Last accessed 9 November 2015].

Panitch, J M and Michalak, S (2005). *The Serials Crisis*. Chapel Hill, NC: University of North Carolina.

Pinfield, S, Salter, J and 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*, early view. DOI: <http://dx.doi.org/10.1002/asi.23446>

Power, M (1999). *The Audit Society: Rituals of Verification*. Oxford: Oxford University Press.

RCUK (2013). *RCUK Policy on Open Access and Supporting Guidance*. Available at <http://www.rcuk.ac.uk/RCUK-prod/assets/documents/documents/RCUKOpenAccessPolicy.pdf> [Last accessed 2 November 2015].

RCUK (2013a). *RCUK announces block grants for universities to aid drives to open access to research outputs*. Available at <http://www.rcuk.ac.uk/media/news/121108/> [Last accessed 2 November 2015].

RCUK (2014). *RCUK APC fund distribution*. Available at http://www.rcuk.ac.uk/RCUK-prod/assets/documents/documents/RCUK_APCfundDistribution.pdf [Last accessed 3 November 2015].

RCUK (2015). *Government funding*. Available at <http://www.rcuk.ac.uk/about/aboutrcs/governmentfunding/> [Last accessed 2 November 2015].

Reckling, F and Rieck, K (2015). *Austrian Science Fund (FWF) Publication Cost Data 2014*. figshare. DOI: <http://dx.doi.org/10.6084/m9.figshare.1378610>

Research Information Network (2010). *Challenges for academic libraries in difficult economic times*. London: Research Information Network. Available at <http://www.rin.ac.uk/system/files/attachments/Challenges-for-libraries-FINAL-March10.pdf> [Last accessed 8 November 2015].

Shieber, S M (2009). Equity for Open-Access Journal Publishing. *PLOS Biology*, 7(8): e1000165. DOI: <http://dx.doi.org/10.1371/journal.pbio.1000165>

Sikora, A and Geschuhn, K (2015). Management of article processing charges – challenges for libraries. *Insights*, 29(2). DOI: <http://doi.org/10.1629/uksg.229/>

Suber, P (2012). *Open Access*. Cambridge, MA: MIT Press. Available at https://mitpress.mit.edu/sites/default/files/9780262517638_Open_Access_PDF_Version.pdf [Last accessed 28 October 2015].

The Economist (2012). The Price of Information. *The Economist*, 4 February 2012. Available at <http://www.economist.com/node/21545974> [Last accessed 9 November 2015].

Ware, M and Wabe, M (2015). *The STM Report: An overview of scientific and scholarly journal publishing*. 4th ed. The Hague: International Association of Scientific, Technical and Medical Publishers. Available at http://www.stm-assoc.org/2015_02_20_STM_Report_2015.pdf [Last accessed 28 October 2015].

Wellcome Trust (2015) *Annual Review 2014*. London: Wellcome Trust. Available at http://www.wellcome.ac.uk/stellent/groups/corporatesite/@sf_cross_cutting_activities/document_s/web_document/wtp058192.pdf [Last accessed 13 November 2015].

Wellcome Trust (2015a). The Reckoning: An Analysis of Wellcome Trust Open Access Spend 2013-14. Available at <http://blog.wellcome.ac.uk/2015/03/03/the-reckoning-an-analysis-of-wellcome-trust-open-access-spend-2013-14/> [Last accessed 28 October 2015].

Wexler, E (2015). What Open-Access Publishing Actually Costs. *Chronicle of Higher Education*, 9 November 2015. Available at <http://chronicle.com/article/What-Open-Access-Publishing/234108> [Last accessed 9 November 2015].

Wijkhuijs, J (2015). Dutch universities start their Elsevier boycott plan. *Univers*. Available at <https://universonline.nl/2015/07/02/dutch-universities-start-their-elsevier-boycott-plan> [Last accessed 9 November 2015].

Wilson, M C (2014). Market failure in the research world. *Public Address*. Available at <http://publicaddress.net/9549> [Last accessed 7 November 2015].

Worthy, B (2013). *David Cameron's Transparency Revolution? The Impact of Open Data in the UK*. Available at SSRN: <http://dx.doi.org/10.2139/ssrn.2361428>

Young, P (2009). *The Serials Crisis and Open Access: A White Paper for the Virginia Tech Commission on Research*. Virginia Tech. Available at <https://vtechworks.lib.vt.edu/handle/10919/11317> [Last accessed 9 November 2015].