Bibliodiversity in Practice: Developing Community-Owned, Open Infrastructures to Unleash Open Access Publishing
Lucy Barnes, Rupert Gatti

To cite this version:

HAL Id: hal-02175276
https://hal.archives-ouvertes.fr/hal-02175276
Submitted on 5 Jul 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Academic publishing is changing. The drive towards open access (OA) publishing, which is being powered in the UK by funding bodies (SHERPA Juliet), the requirements of the REFs 2021 and 2027 (UKRI, Hill 2018) and the international movements such as PlanS ('About Plans'), has the potential to completely transform the established ways of publishing academic research – and we are witnessing substantial changes within the research journal ecosystem. Book publishing, however, is not moving quite so quickly. While funders are clearly signalling a direction of travel for monographs towards OA, most are at present making exceptions for monographs in their OA policies.

In this article we look at the main features of the existing monograph publication and distribution ecosystem, and question the suitability of this for open access monographs. We look specifically at some of the key economic characteristics of the monograph publishing market and consider their implications for new infrastructures designed specifically to support open access titles. The key observations are that the production of monographs displays constant returns to scale, and so can (and does) support large numbers of publishing initiatives; at the same time the distribution and discovery systems for monographs display increasing returns to scale and so naturally leads to the emergence of a few large providers. We argue that in order to protect the diversity of players and outputs within the monograph publishing industry in the transition to open access it is important to create open and community-managed infrastructures and revenue flows that both cater for different business models and production workflows and are resistant to take over or control by a single (or small number) of players.

The scholarly monographs industry

Monographs remain a very important publishing option for scholars, particularly in Humanities and Social Sciences (HSS) disciplines. This has been highlighted by numerous
commentators (e.g. Crossick 2015), and can be clearly seen in the submission data for books and monographs to the last REF assessment conducted in the UK, REF2014.

Figure 1: Books and book chapter submissions to REF2014.

<table>
<thead>
<tr>
<th>REF2014 submissions</th>
<th>All panels</th>
<th>Arts &amp; Humanities (Panel D)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Authored books</td>
<td>12,873</td>
<td>6,477</td>
</tr>
<tr>
<td>Edited books</td>
<td>2,066</td>
<td>1,688</td>
</tr>
<tr>
<td>Chapter in books</td>
<td>13,253</td>
<td>10,003</td>
</tr>
<tr>
<td>Scholarly editions</td>
<td>436</td>
<td>348</td>
</tr>
<tr>
<td><strong>Totals</strong></td>
<td><strong>28,628</strong></td>
<td><strong>18,516</strong></td>
</tr>
</tbody>
</table>

Source: Tanner 2016, Fig. 1, p. 13

Simon Tanner (2016) studies these book submissions to the REF2014 in greater detail, and highlights the remarkable range and diversity of publishers involved, many of which are small and/or specialist presses. He notes that in total 1,180 unique publishers were associated with the books submitted to Panel D alone (Arts and Humanities), with the top 10 publishers accounting for less than 50% of submissions.

Figure 2: The most submitted publishers in REF2014, Panel D (capped at 20 books or more submitted).

Source: Tanner 2016, Fig. 4, p. 18

Remarkably, the vast majority of works submitted were by publishers based in the UK – and I would expect to find similar local bias in other countries, especially within Europe where linguistic diversity encourages regional specialisation.
Similarly, within OA book publishing we have witnessed the emergence of number of new OA initiatives in recent years – primarily spearheaded by small scale academic-led initiatives. A JISC study (Adema and Stone 2017) of new university and academic-led open access publishers in the UK reported that 3 new university presses and 13 academic-led presses had been created in the previous 10 years. The motivations and funding models for each were all slightly different (which is important to note in its own right), but typically they revolved around opportunities to engage with new technologies, new audiences, or subject areas that were not catered for by traditional publishers.

Diversity of publishers is hugely important to the monograph ecosystem – it provides for a diversity of specialisation, editorial positions, publishing objectives and (as we have witnessed with OA publishing) a fertile bed for innovation in publishing practices and outputs. Clearly, in transitioning to an OA publishing world it is important that we protect, sustain and facilitate the diversity and dynamism of the industry as a whole (as opposed to necessarily protecting the existing players). So – how can we create an ecosystem to sustain this? How do we ensure that there are viable revenue flows, broadly accessible technologies to reduce costs, and low barriers for entry by new initiatives?

**Production costs**

Long tails in the distribution of producers, as we see in book publishing, are fairly typically associated with industries characterised by constant or increasing returns to scale. ‘Small’ remains an economically viable size for a publisher. A recent study of US University Presses (Maron et al. 2016) estimated the costs of producing a research monograph and showed that, while the diversity of costs was large, costs per title do not appear to fall as the size of the publisher increases (see Table 1). Of course we don’t know if the bigger publishers have the highest costs because they publish more complex books! But as the lowest cost title for the largest publishers was very close to the median cost recorded for the smallest presses, this data suggests that there is no significant advantage to being large as far as book production is concerned.

<table>
<thead>
<tr>
<th>Publisher Size</th>
<th>Average</th>
<th>Median</th>
<th>95th Percentile</th>
<th>5th Percentile</th>
<th>Highest Cost Title</th>
<th>Lowest Cost Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>Small</td>
<td>$30,091</td>
<td>$27,955</td>
<td>$57,991</td>
<td>$18,678</td>
<td>$65,921</td>
<td>$16,401</td>
</tr>
<tr>
<td>Medium</td>
<td>$44,906</td>
<td>$42,851</td>
<td>$69,417</td>
<td>$26,292</td>
<td>$129,909</td>
<td>$19,516</td>
</tr>
<tr>
<td>Large</td>
<td>$34,098</td>
<td>$33,199</td>
<td>$53,084</td>
<td>$18,149</td>
<td>$76,537</td>
<td>$15,140</td>
</tr>
<tr>
<td>V. Large</td>
<td>$49,155</td>
<td>$48,547</td>
<td>$73,885</td>
<td>$31,760</td>
<td>$99,144</td>
<td>$24,234</td>
</tr>
</tbody>
</table>

A similar study of Dutch publishers (Ferwerda et al. 2013) estimated production costs to be in the range of €5,000 – €10,000, i.e. significantly lower than those reported for the US. These figures are much closer to the experience at Open Book Publishers (OBP)\(^1\).
Following the methodology from the Maron et al. study as closely as possible, Gatti and Mierowsky (2016) report average book production costs at OBP of £5,555 ($8,333).

It remains an open question why the costs reported by Ferwerda et al. and at OBP are so much lower than those reported for traditional US university presses – but I expect one reason for the lower costs is OBP’s ability to start with a born-digital workflow. My own impression is that few (if any) of the traditional monograph publishers presently embrace IT as part and parcel of their workflow and mission in the same way the large textbook and journal publishers do.

**Distribution and revenue flows**

Of course production costs are only one part of the publishing equation. Once published, books need to be discovered by, and accessible to, readers. And publishers need to access a revenue stream in order to cover the production costs.

The distribution of both printed and e-book editions is dominated by a small number of large companies (differing, with very little overlap, between printed and digital content) specialising in either library sales or public sales – indicative of an industry with increasing returns to scale.

Sales to individual consumers, and print sales to libraries, typically result in a copy of the book traveling from the publisher to the consumer via the distributor. In addition, many libraries now rely on distributors to actually select the titles to purchase for their collections on their behalf (for both print and digital editions), and for hosting (and controlling access to) any digital content sold to libraries.

Revenue flows work in the reverse direction – moving from the final customer to the retailer, through to the distributor, the publisher, and finally to the author – with each taking their own cut along the way. It is not uncommon for the retailer and distributor to collectively take a share of at least 45% of the sale price of the book, while the share flowing all the way through to the author is typically less than 5%.

One aspect of this distribution system that is relatively efficient is that customers, retailers and publishers are required to establish and manage order and payment processes with only a small number of distributors – reducing significantly the number of relationships that need to be maintained. If we have $m$ retailers ordering books directly from $n$ publishers then collectively $m \times n$ channels are required to be maintained, while if both sides are operating through a single distributor the total number to channels required falls to $m+n$ (with each retailer and publisher requiring just one) a substantial saving if $m$ and $n$ are large.

For open access books in the digital age almost everything is wrong with this system however. To start with, the costs involved for both parties in establishing direct customer to publisher links have fallen considerably. The majority of print work sales at OBP, for example, are made by customers directly engaging with OBP (usually via our website) rather than through intermediaries. While the need to transfer printed works from the publisher to the customer remains, there is no longer a specific need to do this through an (expensive) third party distributor. Using Print on Demand technology we instruct the books to be printed for each order, and the printer dispatches the books directly to the customer using either the postal system or a courier service. In consequence our total
distribution costs are less than 10% of traditional processes for printed works, and of course almost zero for digital content.

For open access e-books the existing system is even worse: when revenue is based on a sales commission most distributors will not accept a zero price for content, and although there is no need for them to restrict access rights when the work is open access it is technically difficult for distributors to release control over access for only a subset of content and they have little incentive to do this if that subset generates no revenue. Libraries, meanwhile, have no facility to host this content themselves, and also rely on distributors to provide title metadata to ingest into their catalogues. Alternative distribution networks for open access content will need to include new hosting solutions, processes for the provision of metadata, and processes for libraries to make selection decisions.

Certainly there is an opportunity for a wholesale reform of the existing distribution systems for both printed and digital works. However, the underlying feature that distribution networks, especially digital distribution systems, display scaling advantages is likely to remain – and herein lies a significant challenge to the scholarly community. If distribution is ‘naturally’ going to be dominated by one (or a small number) of large enterprises, then these ‘platforms’ will be in a position to control the nature of content distributed, and restrict access to their channels to extract ‘economic rent’ from their centralised position. Such centralization will in its turn endanger bibliodiversity, cost efficiency, product innovation, and the dynamic entry of new publishers.

For commercially operated journals the default revenue stream for open access content has become an author-pays model, with journals imposing article processing charges (APCs). It is worth noting that many OA journals (typically non-profit journals) operate without APCs – typically relying on (often implicit) support from scholarly communities or institutions, and the availability of open source software (such as PKP’s Open Journal Systems) to produce the journal and publish the articles.

The APC funding model reverse the typical financial flows through the system – starting with the author and then progressing to the journal. The experience of wholesale APC funding in the UK demonstrates that actually implementing this ‘change in direction’ has been a costly administrative process for universities to implement. But an important feature of this funding process is that the revenue is tied with the publication – so the only means of providing revenue to intermediate services such as infrastructure lies with a payment per publication. Direct institutional funding for infrastructure to support OA dissemination is not possible within this funding model.

To the extent that sales revenue from print editions will not be sufficient to offset the initial production costs for open access books, new revenue streams need to be developed and, following journals, book processing charges (BPCs) have emerged as the default solution for the majority of ‘legacy’ book publishers. BPCs of over £10,000 per title are not uncommon amongst traditional publishers (Ferwerda et al. 2013), raising serious cultural, ethical and financial difficulties, particularly in Humanities and Social Science disciplines (Eve 2014). We initially will concentrate exclusively on the financial considerations if BPCs were broadly adopted as the primary financing model for open access books. As a starting point, Eve et al. (2016) consider the costs that might have been incurred by paying BPCs for book submissions in the UK to the last REF: taking a BPC of £7,500, and assuming that (due to various exemptions proposed) this is paid on 75% of the books submitted, they estimate that it would have cost “… approximately £96m investment over
the census period. This is equivalent to £19.2m per year. Academic library budgets as they are currently apportioned would not support this cost. However, these sums are but a fraction of the total quality-related funding, Arts and Humanities Research Council and Economic and Social Research Council budgets.” (p.1)

22 As the authors note, whether £19.2m per year is a big number or not depends on your reference point but, in comparison to the maximum allowed annual UUK OA block grant funding of £24 million\(^1\), the figure is high. We should also bear in mind that not all books (e.g. edited or critical editions) are likely to have been submitted to REF2014, so the numbers involved are likely to be larger than that still. UUK are still considering their policies for OA books, but I suspect that BPCs as high as this mean the overall costs involved are still too large to be realistically met by present funding capacity within the UK.

23 But BPCs are not the only way to finance open access books. It is notable that very few of the newly founded OA presses rely on BPCs to finance their operations (Adema and Stone 2017). Many of the new University Presses are seen as an important branding opportunity for the University as a whole, and so are being directly subsidised. Harnessing the financial support of library collectives is another increasingly common financing strategy for OA publishing – particularly in the US.\(^4\) We are also witnessing publishers, such as OpenEdition and the OECD Press, introducing ‘Freemium’ models – where the basic content is provided for free but payment is made for ‘premium’ services provided either to the reader or an institution.

24 At Open Book Publishers we have adopted a mixed funding approach. OBP does not charge authors anything to publish (so, no BPC) but we do encourage them to apply for any publication grants they are aware of. About one third of our funding comes from such grants – with the remainder coming (in almost equal proportion) from the sales of printed works and through support from a collective of libraries, through our library membership programme.\(^5\)

25 What is increasingly apparent is that a whole raft of alternative non-BPC funding channels are emerging – almost entirely being developed by new entrants to the book publishing landscape, highlighting the importance and power of innovation by newcomers. As Geoffrey Crossick notes in his report on open access monograph publishing (Crossick 2015), a new environment for OA books will need a variety of business models.

26 While new OA presses are effectively creating new business model to sustain the production of books, many are facing the same problems in developing effective distribution processes and disperse revenue channels for their content, and in creating and distributing the metadata required by existing platforms. As none has the financial resources needed to solve all the issues independently a shared infrastructure to implement effective change will be of benefit to all.

27 It matters how this infrastructure is provided, however, and by whom. Geoffrey Bilder, Jennifer Lin, and Cameron Neylon warn “[e]verything we have gained by opening content and data will be under threat if we allow the enclosure of scholarly infrastructures” (Bilder et al. 2015). These concerns are echoed in a recent report by SPARC_NA (Aspesi et al. 2019) which documents the changing business models of the major journal and textbook publishers and their strategic manoeuvring to capture users and usage data across the entire research cycle.
The critical issue is: how do we create infrastructures that will effectively scale across a large number of small publishers, without creating technical and/or financial bottle necks or opportunities for single entities to gain ‘control’ over the entire publishing landscape?

Open and community managed infrastructures: Scaling small

We are seeing an increased recognition of the need to create open, community controlled infrastructures to support open access publishing internationally. Within Europe we have seen the emergence of OPERAS, and within that HIRMEOS, each collaborations amongst open access publishers and other institutions across Europe committed to creating shared and open source book publishing services. More broadly, we have seen calls for the development of collective open access infrastructure emerging from the University of California, and with the recent establishment of the Invest in Open Infrastructure initiative.

OBP has been partner with HIRMEOS and OPERAS, and has recently helped form ScholarLed, an international consortium of five academic-led, not-for-profit, OA book publishers collectively developing open source and community controlled publishing solutions, and to directly implement these within their own workflows.

Collectively the ScholarLed presses have now published over 500 books, and expect to publish over 80 new titles in the coming year. One of their strengths as a collective is that they all have different business models and publishing practices – but none charge BPCs. They also have in common a desire to build an inclusive, open and community managed infrastructure for the effective distribution and discovery of their content. The overriding objective of the collective is to work collaboratively with all actors in the scholarly publishing ecosystem (scholars, libraries, publishers) to create systems that allow diverse, small scale scholarly publishing initiatives to effectively operate, and so to create a robust, inclusive and community managed publishing ecosystem: Scaling Small.

Presently ScholarLed is working in collaboration with the Coko Foundation to create open source software to facilitate production workflows. Working collaboratively with several open access journal publishers (eLife, Hindawi and Europe PMC) the Coko Foundation has been developing a modular, open source framework called PubSweet, based on the React JavaScript library, to update and replace the partners existing production workflow processes. The advantage of the modular design of this system is that publishers can integrate, develop or refine components of the system as they wish – rather than being reliant on a single complete and controlled system, and many of these modules are now operational within the respective publishing platforms. More recently, and in partnership with the California University Press, Coko began extending the project to create a book publishing platform (Editoria) using the same PubSweet framework. Members of ScholarLed are now working directly with the community of PubSweet developers in both the development of Editoria and the creation of modules that can be integrated into both journal and book publishing workflows, such as modules for uploading and managing third party digital material (e.g. images and audio/visual files).

ScholarLed has also spearheaded the formation of a major strategic international partnership aiming to realign OA book publishing discovery and distribution processes.
away from competing commercial service providers to a more horizontal and cooperative knowledge-sharing approach. In collaboration with a number of other institutions, universities and university libraries (including Coventry, Birkbeck, Leister and Durham Universities in the UK, the University of California, Santa Barbara, the British Library, JISC and OAPEN) we are aiming to develop open, community managed infrastructures to facilitate post-publication discovery, distribution and revenue management processes. The funder-acronymed COPIM project (Community-led Open Publication Infrastructures for Monographs) is designed to address the key technological, structural, and organisational hurdles that are standing in the way of the wider adoption and impact of OA books, particularly around funding, dissemination, discovery, reuse, and archiving.

The open and collaborative approach adopted aims to create infrastructure that will enable resilience (Adema and Moore 2018) and diversity among publishers of varying size and experience, and with different objectives and outputs. It makes space for the expertise of OA publishers who have developed innovative, sustainable, and ethical approaches to OA publishing that are responsive to the scholarly community (Bargheer and Pabst 2016), and it is flexible enough to allow for a variety of business models. The intention is therefore that OA book publishing will grow through a proliferation of diverse organisations working in different ways, rather than one or two OA publishers growing in size, or a small number of dominant commercial publishers or platforms establishing a restricted (and potentially costly) set of workflows, funding pathways and business models.

Specifically, the objectives of this project are to create:

i. Revenue platform: A modular, scalable revenue generation and management platform for OA books, to be made available to publishers and libraries. Working with a team of publishers, knowledge managers, and librarians the aim is to create the technical infrastructures, organisational processes, financial management procedures, and legal standards to enable new funding channels for financing open access publications to be developed and become sustainable long term.

ii. Business model knowledge exchange: Recognising that the circumstances and opportunities for developing revenue streams and production workflows differ significantly between publishers, we will be working directly with relevant stakeholders from across the sector, including existing scholarly publishers looking to transition to open access, (New) University Presses, library-presses, scholar-led initiatives, academics, learned societies, and open technology developers, to develop sustainable business models. We will collect case studies and associated cost data of different sales and production processes and review the business models of project partners and other publishers in order to identify cost reductions for OA book publishers (e.g. LongLeaf model, infrastructural cost-reductions) of all scales, and will develop an online, open source toolkit for booting-up and running an OA book press. Of course many of the models developed may also benefit from the revenue platform being developed simultaneously.

iii. Open and community controlled governance. Creating durable organisational structures for the coordination, governance and administrative support of the project’s community-owned infrastructure, creating genuine community involvement and collective control (ensuring the infrastructures won’t be governed by a particular commercial interest) is critically important.

iv. Experimental publishing methods and content use: Recognising that open access is potentially much more than just ‘free to read’ content, we will examine ways to more closely align existing software, technologies, workflows and infrastructures for experimental
publishing with the workflow of OA book publishers and with the infrastructures that COPIM will create.

v. Open dissemination and discovery infrastructure: Developing technical protocols and collectively controlled infrastructure to better integrate OA books into institutional libraries, repositories and digital learning systems, while ensuring the infrastructure is resistant to 'capture' by any single entity (either commercial or not).

vi. Archiving and preservation: Ensuring scholarly works can be preserved and remain accessible to future generations is an important part of any effective scholarly publishing process. Existing archiving process are not well adapted to open access and experimental publications and we will be seeking to identify the key challenges associated with archiving open access research monographs in all their variation and complexity, and to develop new solutions that can be implemented by small presses.

Conclusion

Open access publishing processes are at a critical juncture. Within the spheres of both journal and textbook publishing large commercial players are moving to dominate the digital publishing landscape – and with it control the revenue streams and technological development of open access publishing processes. Such dominant players have not yet emerged within the open access monograph space – so there is a window of opportunity to learn from the experience of other sectors and to develop effective infrastructure to sustain and facilitate the bibliodiversity so important to Humanities and Social Science scholarship.

In order to do so we need to create robust, open and community managed infrastructures to support the diversity of open access publishing initiatives. Many commentators have noted the variety of business models and production practices adopted by monograph publishers (particularly new open access publishers) – but quite often this observation is followed by the concern that “none of these have been shown to scale”. We suggest that we should not be looking for a single model to scale and “rule them all”, but to embrace this diversity as an inherent strength of the system as a whole and look to collectively develop infrastructures to support that diversity – that is, to “Scale Small”.

To do so we need to work collectively, across all sectors of the scholarly community. But that, in itself, will not be enough: we also need to actively develop systems resistant to capture and control by any single entity (be they commercial or non-commercial). To do that means creating open, community controlled and interoperable systems, protocols and standards – as well as new collectively managed revenue streams to support open access publishing and distribution processes and infrastructure.
BIBLIOGRAPHY


Academic publishing is changing. The drive towards open access publishing, which is being powered in the UK by funding bodies (SHERPA Juliet), the requirements of REFs 2021 (UKRI) and 2027 (Hill 2018), and Europe-wide movements such as the recently-announced Plan S (‘About Plan S’), has the potential to shake up established ways of publishing academic research. Within book publishing, the traditional print formats and the conventional ways of disseminating research, which are protected and promoted by a small number of powerful incumbents, are being challenged. Academic publishing, and academic book publishing, is at a crossroads: will it find ways to accommodate open access distribution within its existing structures? Or will new
In this article we look at the main features of the existing monograph publication and distribution ecosystem, and question the suitability of this for open access monographs. We look specifically at some of the key economic characteristics of the monograph publishing market and consider their implications for new infrastructures designed specifically to support open access titles. The key observations are that the production of monographs displays constant returns to scale, and so can (and does) support large numbers of publishing initiatives; at the same time the distribution and discovery systems for monographs display increasing returns to scale and so naturally leads to the emergence of a few large providers. We argue that in order to protect the diversity of players and outputs within the monograph publishing industry in the transition to open access it is important to create open and community-managed infrastructures and revenue flows that both cater for different business models and production workflows and are resistant to take over or control by a single (or small number) of players.

AUTHORS

LUCY BARNES
Open Book Publishers
lucy@openbookpublishers.com
(corresponding author)

RUPERT GATTI
Open Book Publishers and Trinity College, Cambridge