



RESEARCH INFORMATION NETWORK

ACTIVITIES, COSTS AND FUNDING FLOWS IN THE SCHOLARLY COMMUNICATIONS SYSTEM IN THE UK

May 2008

Model Structure (Annex B)

Prepared by:

Cambridge Economic Policy Associates Ltd

 **CAMBRIDGE ECONOMIC
POLICY ASSOCIATES**

CONTENTS

1. Introduction	3
1.1. Purpose of paper.....	3
1.2. Methodology.....	3
1.3. Structure of paper	3
2. Value chain and definitions	5
2.1. Value chain.....	5
2.2. Cost concepts & definitions	7
3. Delivery of model requirements.....	8
3.1. Requirement 1: current, system-wide costs.....	8
3.2. Requirement 2: apportionment of costs	8
3.3. Requirement 3: sources and volumes of funding.....	10
4. Model structure Overview.....	11
4.1. Module 1: Costs of publication and distribution (global)	13
4.2. Module 2: Costs of access to journals (UK)	18
4.3. Module 3: Funding and geographic cost allocation module.....	20
4.4. Model calibration	23
4.5. Scenario testing.....	24
Appendix 1: Business model value chains.....	25
A1.1 Introduction.....	25
A1.2 Subscription-based print journal model	26
A1.3 Subscription-based electronic journal model.....	27
A1.4 Open access journal model.....	28
Appendix 2: Cost definitions	29
A2.1 Fixed & variable costs	29
A2.2 Direct & indirect costs	29
A2.3 Average cost (break-even price).....	30

1. INTRODUCTION

1.1. Purpose of paper

This paper presents for discussion our approach to modelling the costs, activities and funding flows involved in the scholarly communications process. The purpose of this paper is to present the scope of the scholarly communications process value chain ('the value chain') activities and outputs modelled by Cambridge Economic Policy Associates (CEPA) in consultation with industry representatives during 2007, and the structure of the spreadsheet model itself.

1.2. Methodology

CEPA spent time at the inception of the project talking to Expert Panel members and other representatives from the research communications industry about the structure and components of the scholarly communications process value chain. We combined the findings from interviews with information from around 20 pieces of literature that we acquired and reviewed (many recommended by interviewees). The principle aim of this process was to define the interdependencies, activities, and actors that make up the value chain.

For ease of reference we reproduce here the model requirements as set out in the terms of reference:

- To produce estimates of the current costs, from a system-wide perspective, of the scholarly communications process as a whole (requirement 1).
- To identify the cash and non-cash costs incurred by agents involved in the various activities and sub-activities of the scholarly communications process in the UK (requirement 2).
- To identify the sources and volumes of funding and other resources made available to meet those costs (requirement 3).

1.3. Structure of paper

The rest of the paper is structured as follows:

- Section 2 presents our understanding of the scholarly research value chain and outlines the broad scope of the modelling to be undertaken;
- Section 3 presents more detail about the key considerations arising from the modelling requirements set out in the terms of reference; and

- Section 4 presents the model structure and details of the overall approach to capturing the activities, costs and funding flows of the scholarly communications process.

Appendix 1 contains our interpretation of the value chain activities, participants and funders by business model. Appendix 2 provides some key cost definitions which feature throughout the document.

2. VALUE CHAIN AND DEFINITIONS

This section discusses the value chain and other key definitions that were required at the outset of the modelling work.

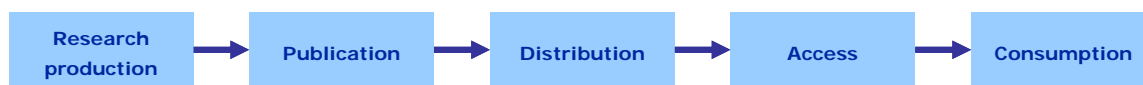
2.1. Value chain

In developing the model for this project, we distinguished the broad ‘research value chain’ from the more focused ‘scholarly communications value chain’. There are a variety of ways in which scholarly information can be distributed prior to publication, including by informal correspondence, conference papers and presentations. However, while this so called pre-publication ‘grey literature’ serves an important role in the communication of research findings, our focus in this work was on the *published research article*.

2.1.1. Research value chain

Figure 2.1 presents what we considered to be the five components that are common to any *research value chain*, regardless of the business model of delivery.

Figure 2.1: Research value chain



The generic components of the research value chain are as follows:

- *Research production* captures the set of activities carried out by researchers to create a research article and to subsequently submit it for peer review and publication. As discussed above, there are different ways of distributing scholarly communication prior to publication.
- *Publication* involves the peer review and editing of articles received, the composition of approved articles into journals, and the branding of the journal.
- *Distribution* covers the logistical activities required to (either physically or electronically transport) copies of journals to libraries and other users.
- *Access* captures the set of activities carried out mainly by publishers and libraries with the purpose of making journals accessible by end users (both now and in perpetuity). It also includes activities undertaken by the researcher in identifying, searching for and accessing articles, mainly online and/or in libraries.
- *Consumption* captures the reading activities undertaken by researchers.

We considered the above five components to comprise the research value chain. The majority of the research activities (and therefore costs) are incurred in the first component – research production.

2.1.2. Scholarly communications value chain

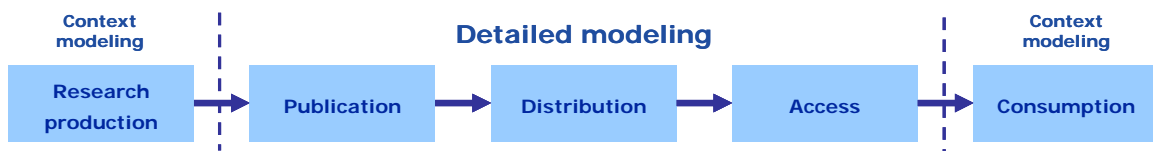
Our definition of the *scholarly communications value chain* comprised the middle three components of the broader research value chain presented above, namely publication, distribution, and access.

Figure 2.2 below presents the components of the scholarly communications value chain in the context of the research value chain, and the extent to which each has been modelled.

Modelling has been carried out on two levels:

- Most of the **detailed modelling** of this assignment was carried out on the publication, distribution and access components (i.e. the scholarly communications value chain), and is captured in a single ‘main model’. As discussed in more detail below, the activities of these components were modelled in terms of volumes of articles flowing through the scholarly communications process (number of articles by subject, journal type etc), and the unit costs (fixed, variable, direct and indirect) per article. Funding flows have been modelled separately.
- As part of the work we also modelled the research production and consumption elements of the value chain. However, this was carried out at a significantly lower level of detail but was considered sufficient to put the scholarly communications costs and funding flows into the wider context of the costs associated with these other elements of the research value chain. We referred to this as **context modelling**.

Figure 2.2: Model scope



For the purpose of the detailed modelling, we have defined the scholarly communications value chain as being in relation to articles that are peer reviewed, where ‘peer review’ is defined as the process of subjecting an author’s unpublished research to the scrutiny of other experts in the same field. The model captures the delivery of peer-reviewed articles through print and electronic journals. Monographs, were excluded from our modelling work.

2.2. Cost concepts & definitions

Appendix 2 sets out a number of cost definitions that are relevant to the structure of the model. We used definitions that are broadly consistent with those suggested in King (2007)¹ – particularly in relation to fixed, variable, direct and indirect costs.

It was considered to be important to distinguish between costs incurred in the scholarly communication process (which includes all aspects of publication, distribution and access) and the prices that libraries (and others) incur in acquiring journals (in print and / or electronic form). We used the following definitions:

- The journal / article **‘acquisition price’** is the price paid by UK libraries and other organisations (special libraries, corporate users, etc) buying journals / articles.
- **‘Incurred costs’** are the total costs of publishing, distribution and access activities incurred by value chain participants, excluding the acquisition price most notably, publishers and subscription agents.

A clear distinction between these two concepts was essential, in order to avoid ‘double counting’ in the calculation of the cost of the scholarly communications process. Incurred costs can be either **‘cash’ or ‘non-cash’ costs** - for the purposes of modelling we defined the latter as costs incurred (typically by academics) which are not directly remunerated as part of the scholarly communications cost.

¹ Donald W King, ‘The cost of journal publishing: a literature review and commentary’, *Learned Publishing*, 20 (2), pp85-106 (April 2007)

3. DELIVERY OF MODEL REQUIREMENTS

In this section we seek to expand on the RIN modelling requirements, to identify the required level of modelling disaggregation, and the key policy questions that the model sought to address. The discussion reflects our analysis of the value chain and different business models, the literature review and the main conclusions of the first Expert Panel meeting.

3.1. Requirement 1: current, system-wide costs

The RIN terms of reference required CEPA ‘to produce estimates of the current costs, from a system-wide perspective, of the scholarly communications process as a whole’. In producing these estimates the model allowed for disaggregation of costs of article production by:

- **Type of costs** (direct and indirect, fixed and variable costs, cash and non-cash).
- The **stage in the value chain** (i.e. publication, distribution and access).
- Different **journal types** which vary by the following factors:
 - Business model (for example traditional publisher ‘subscription’ models, and open access ‘author pays’ models).
 - Subject (Science/Technology, Medical, Arts & Humanities).
 - Journal category (Popular Hybrid Journals, Major Discipline Journals, Niche Journals).
 - Article rejection rate.
- Alternative production/delivery **formats** (print only, print and electronic, electronic only).
- Different **publisher types** (Commercial publishers, University Presses, and Society publishers).

3.2. Requirement 2: apportionment of costs

The terms of reference required that the study “identifies the cash and non-cash costs incurred by agents involved in the various activities and sub-activities of the scholarly communications process in the UK”.

From a modelling perspective there are two separate aspects to this: (i) cash and non-cash costs; and (ii) the local (UK) and global aspects of the value chain.

3.2.1. Non-cash costs

The key ‘non-cash’ cost that was separately identified in the model is the input of the academic community in peer review and editing of journals. This was captured as a non-cash *cost* incurred in article publication, and as a non-cash *funding contribution* elsewhere in the model.

The model also captured the non-cash (time) ‘article search’ and ‘reading’ costs incurred by researchers during research production and consumption. The total costs associated with library and individual users accessing scholarly information (article search/print cost) was modelled contextually.

3.2.2. Geographical scope

The geographical scope of the model varied according to the stage in the value chain being considered:

- In particular, publishing and distribution (as well as research production) were regarded as being global activities. The model identified the total incurred costs associated with publishing and distribution (regardless of where the activity occurs).
- The access component however was considered to be local - the modelling focused on the costs incurred in the UK to provide access.

Given this approach, a key policy question related to what non-cash costs are incurred by UK-based researchers in the scholarly communication process (i.e. carrying out peer review and editing activities)? The model compared these costs (as a percentage of the total non-cash incurred costs in the global system) with: (i) the proportion of articles generated in the UK; (ii) the UK’s contribution to the total publishing/distribution cost through academic journal subscriptions, and (iii) measures of ‘consumption’/level of usage by UK users compared to global consumption.

It is important to note that the model does not capture actual publishers in a geographical context. The model allocates ‘notional’ journals to different publisher types (see section 4) in order to account for different tax treatments and sources of funding, but it does not seek to allocate journal publication to specific countries. We recognised that there may be certain perceived advantages for the host countries arising from the location of publishers, but we did not believe that it is appropriate to seek to capture it in this study².

² In any event, we note that the argument that location of a publisher offers particular tax or macro-economic benefits in the long-run. This is because to do so would require us to assume that the publishing industry is more productive than alternative industries that might exist in its absence in any country. In addition, given the potential for transfer pricing by companies operating in global markets, it is not easily possible to allocate activity and value to a single market.

3.3. Requirement 3: sources and volumes of funding

The focus of this requirement is for the model to support analysis of:

- How the funding flows in the current UK system would change under different business models (i.e. author-side vs. subscription based payment models) and if non-cash cost turn into cash cost (e.g. if publishers have to pay academics for their peer review activities). In order to do this, each of the Journal types (defined in the next section) has separate assumptions as to the sources of funding (advertising, academic and other subscription revenues, author side payments, membership fees, etc).
- The proportion of the total incurred cash and non-cash fixed (of first copy) costs³ that are funded by UK institutions (mainly academic institutions). The model estimates the UK's contribution to the non-cash cost incurred during peer review of articles, and the UK's contribution to the publishing / distribution cash cost through libraries' and other journal subscriptions.

³ The exclusion of direct and indirect variable costs reflects that these are expected to be paid by the purchasing institution and are assumed therefore to be incurred UK institutions in proportion to their use / consumption of journals.

4. MODEL STRUCTURE OVERVIEW

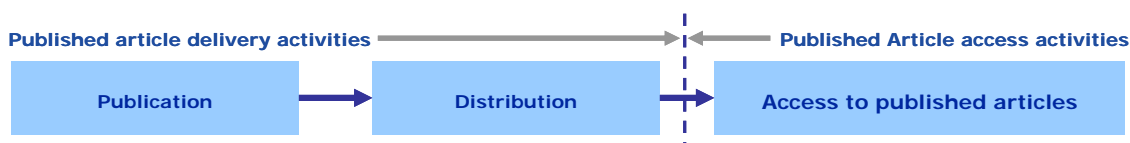
In this section we provide an overview of the model structure.

At the highest level, the model is organised around:

- The grouping of activities in the value chain as illustrated in Figure 4.1, i.e. the delivery of published articles⁴ (including publication and distribution⁵); and access to published articles.
- The two main features of the model requirements, i.e. the incurred costs of the value chain as a whole (requirements 1 and 2) and the sources and volumes of funding.

The model is a single, rather than a multi-year model. This means that exploration of dynamic aspects of the sector requires the user to provide inputs in a separate version of the model. Therefore, scenario analysis is made across separate versions of the model.

Figure 4.1: Article allocation versus article access



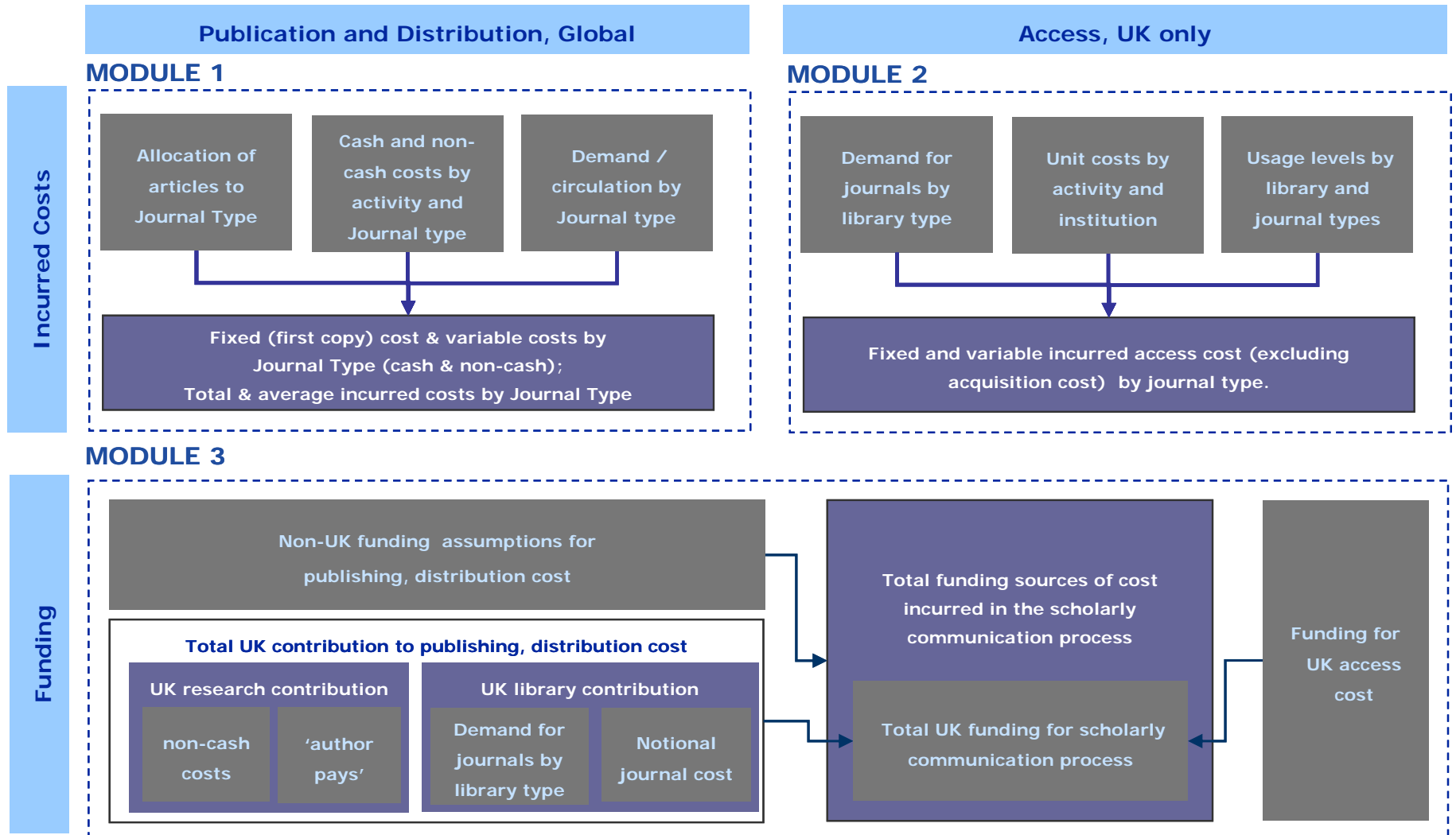
Research production activities which precede the publication component and *research consumption* activities which come after the access component of the value chain, are not modelled in detail. They will be instead modelled at a high level to set the contribution of the above three components in the context of the broader *research* value chain.

The model is divided into three **modules**, each comprising of a number of **input components** generating sets of **outputs**, shown in separate input and output worksheets. Figure 4.2 below provides a summary of the model structure, including the main input components and outputs of each module (in the form of a two by two table). The following sub-sections discuss how the modules interact to deliver the model requirements.

⁴ The model will be structured around flows of *published* articles (and journals) to ensure a degree of definitional precision in our approach.

⁵ We consider the provision of electronic journals through publishers' servers or through online portals (either open access or subscription-based) as simultaneous distribution and access. In order to allow for ease of analysis of costs by UK libraries under the access component, we classify electronic journal provision under the distribution component of the value chain.

Figure 4.2: Scholarly Communications Process Model Modules



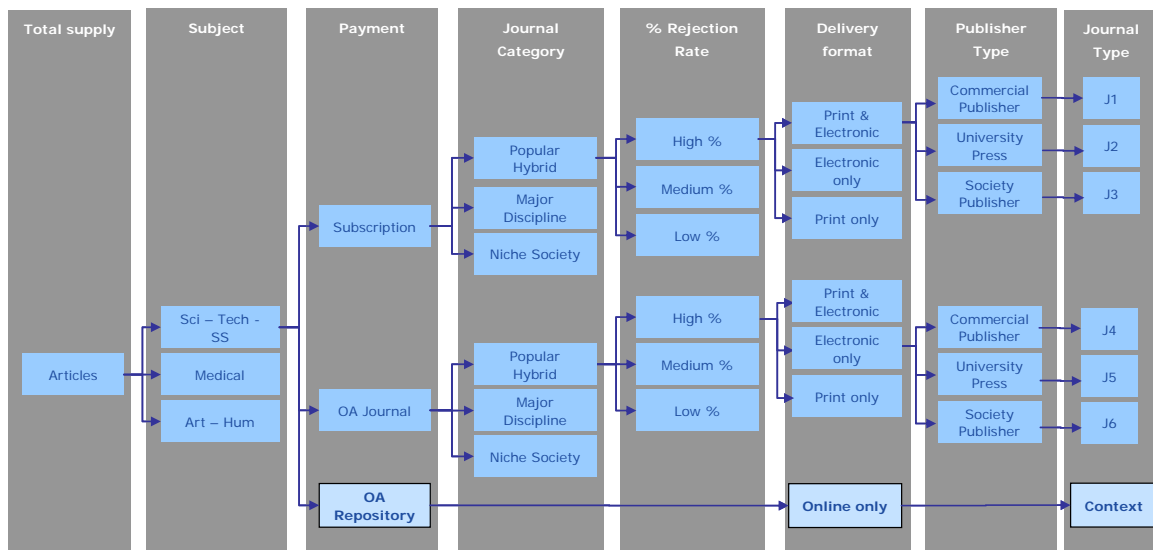
4.1. Module 1: Costs of publication and distribution (global)

Module 1 calculates the costs incurred in carrying out the activities within the publication and distribution components in the scholarly communication process value chain. The global nature of the scholarly communications industry means that this module considers the publication and distribution of all journals. As has already been noted in Section 3 above, the model does not separate cash costs in the UK from those incurred elsewhere, with the exception of the non-cash costs incurred by UK-based academic peer reviewers and editors.

4.1.1. Article allocation component

The **article allocation component** of the model allocates the inputted total volume of research articles entering the system in a year to journal types. Journal types are classified by subject, business model of delivery, journal category, rejection rate, delivery format and publisher type. Figure 4.3 shows how the article allocation component is structured. An assumed supply of articles is allocated by each classification to a series of Journal types. We discuss each of our proposed classifications below.

Figure 4.3: Article allocation component



Subject

The supply of articles is firstly allocated by subject type. We distinguish 3 subject areas: science, technology & social sciences; medical; and arts & humanities. For simplicity and to keep the model tractable, we only sought to model these three classifications. We assumed that there are not significantly different costs associated with journal production *between* specific subjects *within* each classification – other than in relation to editing (typesetting / formatting costs) of articles within subject fields with large numbers of formula or complicated diagrams.

Business Model

Articles are then allocated to either a subscription-based business model or an author-side payment journal (open access journal)⁶.

Journal category

The model then allocates the supply of articles to different journal categories – which have different cost structures (and also revenue / funding sources). For the purpose of modelling simplicity, we aggregated articles into three journal categories⁷:

- **Popular hybrid journals**, examples of which might include titles such as *Science* and the *Journal of the American Medical Association*. Popular hybrid journals are strongly branded, are high circulation titles featuring a range of fairly broad, widely accessible articles and will be funded partly by advertising. Journals in this category might typically be subscribed to by more than 1,000 institutions. Anecdotal estimates indicate that less than 5% of journals fall into this category.
- **Major discipline journals** examples of which might be *Nanotechnology*, the *British Journal of Cardiology*, and the *Journal of Applied Economics*. Major discipline journals are subject-specific, still quite strongly branded with relatively high circulation compared to the niche titles. They will be purchased on a subscription basis by the vast majority of university libraries. Journals in this category are typically subscribed to by 500-1,000 institutions. About 20% of all journals published are considered to fall into this category.
- **Niche journals**, examples of which might include the *Canadian Journal of Botany* and the *Scottish Journal of Theology*. Niche journals are highly specialised with less than 500 annual subscriptions. They are mostly funded by society members. Up to 75% of journals are captured in this category of journal (i.e. the tail in the distribution of journals by circulation).

Quality (rejection rate)

The model also allocates articles by an assumed level of quality or peer review. We assumed that articles (manuscripts) allocated to journals with high rejection rates would undergo more rigorous – and would therefore be more costly in non-cash terms – initial rejection rates and peer review. This also captures rejected articles redirected by editors to journals with a

⁶ For completeness, the model simply captures an open access repository as an additional *journal type*.

⁷ Review journals, examples of which include the *British Journal of Hospital Medicine*, *Alternative Medicine Review* and the *Technology Review Journal* contain articles which have been specifically commissioned by either the publisher or the learned society. The model captures review journals as *Major Discipline* journals.

different focus more suitable to the manuscript contents and/or a lower level of rejection rates.

Delivery format

Articles are also be allocated to a delivery format, namely: print only; print & electronic; or electronic only. The marginal cost of a print journal (or a print & electronic journal) was assumed to be considerably higher than the marginal cost of an electronic only journal. Therefore the model captures how the supply of articles is allocated across these different journal delivery formats.

Publisher type

Finally, articles are allocated to three different **Publisher types** to capture different average rates of return after accounting for tax treatment, and different sources of revenues by journal. We defined three publisher types:

- **Commercial** publishers capture the handful of large players in the scholarly journal industry. These are for-profit, tax paying private sector companies publishing a broad range of titles across subject areas, all typically offering ‘big deal’ packages of titles to libraries.;
- **University press** publishers capture Oxford and Cambridge University Presses, as well as equivalently large organisations in the US. These are not-for-profit organisations although they do exist to generate a surplus which is redistributed back into publishing and/or other activities carried out by the university; and
- **Society** publishers capture the high count of smaller, membership based publishers, and include smaller university presses not captured above (e.g. Liverpool University Press). Publishers in this category redistribute any surplus generated back into the publishing and other activities of the society to which they are affiliated.

Journal type

The final output of this component of module 1 is a long series of journal *types* (not to be confused with journal *categories*), created by the supply of articles allocated according to the classifications set out above. The distribution of articles into *journal types* reflects the input assumptions of the model user.

4.1.2. Activity and journal type unit cost input component

Each journal type features a series of associated unit costs, assigned in the **unit cost component** of module 1⁸.

Unit costs vary by activity. While the set of activities required to publish and distribute each journal type are broadly similar (with some obvious exceptions such as no printing activities required in the publication of electronic journals), the per article and/or per journal unit costs of each activity differ depending on the journal type⁹. It is important to recognise that these are *incurred* costs – i.e. the total costs of publishing and distribution activities incurred by publishers (and distributors).

The combination of unit costs and activities generate a fixed (first copy) cost of each article and, once multiplied with the average number of articles per journal type, fixed (first copy) cost of the journal types. While most of the first copy activity costs are cash costs incurred by publishers, some first copy costs are incurred by peer reviewers and will therefore be *non-cash* costs.

The model also adds the following unit costs per article:

- *Variable publishing & distribution cost.* By definition the total variable costs will be a function of total circulation of the journal category. Variable costs typically include subscription administration, online hosting, printing and inventory management, and delivery/fulfilment cost.
- *Indirect costs* (e.g. marketing and other overheads). Assuming that these costs are fixed per journal implicitly assumes that there are constant returns to scale¹⁰. We did not explore this assumption empirically in the journal publishing sector, but we believed that it was a reasonable simplifying assumption.
- *Publisher's return / surplus.* For simplicity, we assumed that the return required by publishers is an average mark-up by article type on the total fixed and variable cost. We recognised that this was a simplification of the real world in which returns on journals vary according to their maturity, but believed that it was not unreasonable.

4.1.3. Subscription input component

A **subscription component** captures the number of subscriptions of different journal types to aggregate the per journal fixed and variable costs to estimate the average cost (or break-

⁸ The categories of costs (or 'cost types') we describe in this section broadly follow the definitions set out by King (2007). For ease of reference, appendix 3 provides a short explanation of cost types.

⁹ Appendix 1 contains our mapping of the value chain of three key business models.

¹⁰ This is necessary simplification since we do not propose to seek to model actual publishers. It means that the indirect costs (i.e. overheads) associated with a publisher's activities grows in proportion to the number of journals that it publishes.

even price¹¹) incurred in the publication and distribution components of the scholarly communications process, by subject, business model, etc. A count of subscriptions is applied to each journal type and can be adjusted to provide an understanding of changes in the break-even price of each journal type. The model assumes different subscription numbers for academic and other subscribers, and break-even prices differ for subscribers.

We did not make the subscriptions input dynamic – i.e. linked to a demand function (including a specification of prices, and price elasticities).

4.1.4. Module 1 outputs

Policy spreadsheet models tend to feature numerical ‘unit count’ outputs and financial/economic outputs; but are also useful for presenting value chains in a detailed, activity-based manner which may not have been presented previously. In developing this model we distinguished **final**, **intermediate** and **contextual outputs**. The *final* outputs of module 1 include calculations of:

- Fixed (first copy) costs per article by journal type (cash and non-cash);
- Variable costs by journal type ; and
- Total, average and marginal costs of publication and distribution activities per article and journal type.
- Estimates of ‘break-even’ prices by journal type and subscription number.

It is possible to analyse all of the above costs by subject area, business model, journal category, quality (rejection rate), delivery format, and publisher type.

Module 1 also generates *intermediate* outputs that feed in as inputs to the calculation of the financial outputs presented above, and that themselves contribute to the modelling requirements outlined above. Intermediate outputs include counts of articles, and consequently the number of journal types, flowing through the publication and distribution components of the value chain; as well as counts of circulation of different journal types. As noted, these would be expected to be calibrated to give a reasonable approximation of the current supply of journals in scholarly communications process.

Bundling or ‘big-deals’

The model captures costs and funding flows both in the presence and absence of ‘big deals’ in the market. For simplicity, the model calculates break-even prices (average costs) of journal types by picking-up different assumptions about: (i) the number of subscriptions to each journal type; and (ii) the step-change in the total cost of each journal type arising from a change in the notional mark-up charged by different publisher types when offering the big deal. Further to this, the model calculates:

¹¹ See appendix 2 for a short algebraic explanation of average cost.

- **changes in the level of access provision.** The model adjusts the notional library subscription portfolios in the presence of big deals for major discipline and niche journals published by Commercial publishers and University presses;
- **changes in libraries’ (non-subscription) access cost.** The model calculates the costs incurred by increased levels of journal access commensurate with increases in libraries’ portfolios of electronic subscription journals arising from big deals.

4.2. Module 2: Costs of access to journals (UK)

Module 2 captures the volumes of activity carried out, and the associated costs incurred, by different types of university libraries in providing access to journals and articles in the UK¹². The costs incurred by individual users, corporate entities and government in accessing peer reviewed research are captured by special libraries, but are contextually only. We believed that this was reasonable given the focus on the academic sector.

The cost of acquiring journals incurred by libraries was not captured in this input component in order to avoid ‘double counting’ the cost of creating journals is captured in module 1. This cost to libraries is considered as part of module 3 – relating to funding.

4.2.1. Library type component

The *library type* component of module 2 captures the university-based composition of journal access provision in the UK and contextually models journal portfolios of special libraries. The categorisation in this module is used to distinguish different library cost structures; and in other parts of the model to consider the proportion of UK’s funding of the scholarly communications system that are apportioned to the different library types.

The model defines university library types using the Society of College, National and University Libraries’ (SCONUL) university library classification scheme. The SCONUL classification distinguishes four types of university library¹³:

- **CURL:** Members of the Consortium of University Research Libraries. These are the libraries of 25 or so large UK universities plus the British Library, National Libraries of Wales and Scotland, the V&A Museum, and the Wellcome Library for the History and Understanding of Medicine;

¹² We use the term ‘access’ to capture access to new journals and access to archive journals. Therefore, the term captures indexing, archiving and preservation activities generally carried out by libraries.

¹³ A small number of SCONUL members fall outside of this classification such as independent institutions which do not report to the Higher Education Statistics Agency (HESA). For simplicity we propose to exclude this category from the analysis

- **Old:** Those universities founded or chartered before the Education Reform Act 1992, including the Open University, but excluding CURL members;
- **New:** Universities incorporated in 1992 or subsequently. These include the former polytechnics, and some former HE colleges; and
- **HEC:** Higher education colleges which do not have formal university status. This group covers large general colleges (e.g. University College Northampton) to small specialist institutions (e.g. Royal College of Art).

A fifth category of library is modelled contextually: **Special Libraries** include smaller public and corporate libraries.

4.2.2. Journal demand and access components

This component of the model captures: (i) journal demand by library type; and (ii) level of access or ‘usage’ by journal type. The total, average and marginal costs incurred by university libraries in carrying out their journal access provision function vary by:

- The number and types of journal universities subscribe to. For example, a *CURL* university library offers access to a higher number of journals than a *new* university library, due to the relative emphasis placed on research within each institution. The model accounts for different library subscription portfolios (in the presence and absence of big deals) and defines the format of delivery for each journal subscription.
- The number of researchers and students accessing the facilities, and their level of usage. The model captures and applies assumptions about the level of usage for each journal type.

4.2.3. Unit costs input component

Each library type is involved in a broadly similar set of activities in carrying out their journal access provision function. The activities are those set out in the value chains in Appendix 1 and include procurement of journals; indexing and archiving; user account management and search support.

As in module 1, fixed and variable, direct and indirect unit costs are inputted in this component.

- *Fixed journal access costs* are defined as the costs that the library incurs to provide access to a journal regardless of the level of usage. Depending on the format of delivery, these include the cost of subscription processing, journal collection development, negotiations and licensing, cataloguing and linking, and archiving and preservation of journals in the library system. Further costs relate to the provision of physical space for storage of print journals, and reading access. Fixed journal access cost cover the overheads associated with the infrastructure (e.g. IT) and management

of the library. The model accounts for the portfolio of journal types and their format of delivery which significantly influences the cost structure of libraries. (Note: This definition is similar to that in Module 1, where the fixed costs are by journal types)

- *Variable costs* of libraries relate to access by researchers only. These costs include, for example the provision of usernames and password and the on-going administration of user accounts, the provision of server capacity to host username, password and usage information and staff time.

In addition, this component accounts for variable usage cost of research, which include costs incurred in travelling to libraries, search cost and costs of photocopying or printing articles.

4.2.4. Module 2 outputs

The key *final* outputs that module 2 generates are:

- The total incurred fixed costs of providing journal access by different university library types;
- The variable costs (user circulation cost) of providing access to different journal types (where the key differences are between print and electronic format journals);
- The average user cost of article access (cost per use).

4.3. Module 3: Funding and geographic cost allocation module

Modelling requirement 3 and the geographical aspects of requirement 2 are captured in the **Funding allocation module**. The key points of context for this module are: (i) the observation that the points on the value chain where costs are *incurred* are not necessarily the same points as where the costs are *borne*; (ii) only a proportion of the total incurred costs of the system are borne by UK institutions.

The two key objectives of the module are therefore:

- To allocate the total publication and distribution ‘cash’ costs of the scholarly communications process to different sources of funding.
- To identify the proportion of total funding for both ‘cash’ and ‘non-cash’ incurred costs that are borne by UK academic institutions (i.e. ‘geographic funding allocation’).

4.3.1. Funding sources component

This element of the model identifies the proportion of the total publication and distribution costs that are borne by different sources. These sources of funding include:

- author payments;

- subscriptions from academic and public sector libraries;
- subscriptions and payments from the other users and the corporate sector;
- advertising revenue; and
- membership fees (for society publishers).

The funding proportions are specified by journal type. This means that as the model user changes assumptions about the proportion of articles allocated to different journal types, the model generates different funding allocation outcomes. For example if more articles go to journals that have a predominantly ‘author’ pays business, the model shows a shift in the funding flows from subscriptions (e.g. by libraries) to research funding.

4.3.2. UK ‘non-cash’ cost funding component

Module 1 identifies the total ‘non-cash’ costs incurred in the publication and distribution of journals. These costs are calculated based on an estimate of the time input for editing¹⁴ and / or peer-reviewing, which is then applied to an assumption about academics opportunity cost of time¹⁵. This component of the module also allocates the total ‘non-cash’ costs that are attributable to UK academics, compared with academics from the ‘rest of the world’.

Note that, by definition, we assume that the most of the costs incurred by academics in peer review / editorial are borne by academic research funders (in addition to the cash cost incurred by some publishers to pay for peer review / editorial services during the publication of particular journals – these cash costs can include for example, the cost of the administration of online peer review / editing).

4.3.3. UK ‘cash’ cost funding component

This element of the module focuses in particular on the funding of the cash costs of journal publishing and distribution and access borne by UK academic institutions. In practice, because all of the access costs developed in module 2 are UK related costs (i.e. UK library costs) the focus of the module is in seeking to allocate the publishing and distribution costs.

Figure 4.2 below shows in a simplified form how the module does this. It involves the allocation of total publishing and distributing costs according to:

- the different providers of funding (at a global level);
- the proportion of the funding type that is accounted for by the UK. The model calculates: (i) the proportion of total journal cost (by journal type) funded through

¹⁴ Of all editing activities only a part is “non-cash” cost, a major part of editing, particularly among commercial publishers is “cash-cost” and incurred during in-house editing of the articles and the management of the editing process.

¹⁵ There are a range of possible approaches that we would expect to use to identify the appropriate opportunity costs. We do not propose to discuss these issues here.

UK-library subscriptions¹⁶; and (ii) the proportion of ‘author pay’ articles that are produced by UK-funded academics. For simplicity, we assumed that UK located academics are also UK funded.; and

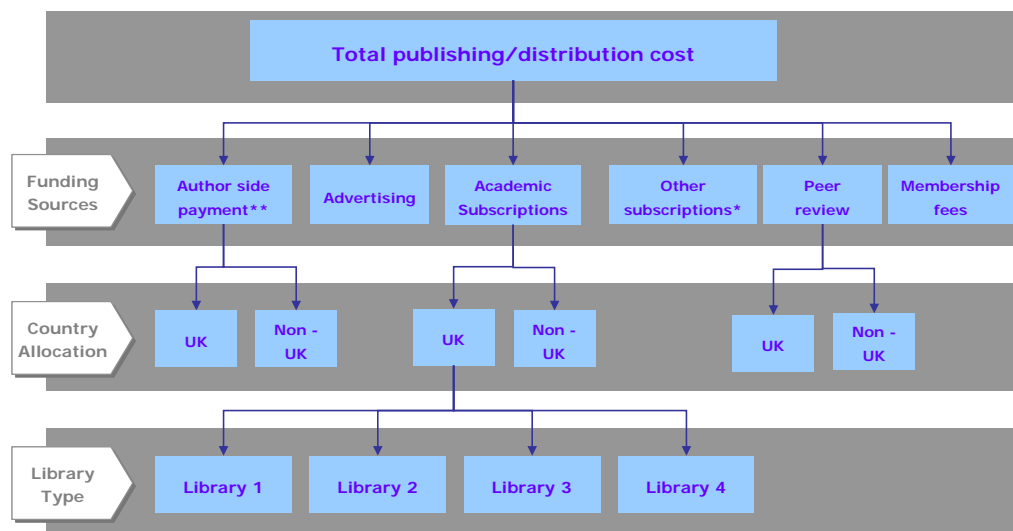
- for the UK funded subscription costs, an allocation to the different library types (discussed in section 4.2.1 above) on the basis of estimates of the proportion of total UK journal expenditure incurred by each library grouping¹⁷.

Although Figure 4.2 shows (implicitly) the total costs of ‘global’ production and distribution, the model is structured to allow the user to input different allocation assumptions for the journal type. For example, an open access journal will have a different ‘funding source’ allocation than a subscription (‘user-pay’) model.

Using these proportions for each journal type it is then possible to provide a ‘top’ down estimate of the costs borne by UK libraries.

In addition to calculating a ‘top-down’ estimate of total library costs we explored the possibility of producing bottom-up estimates that can be compared with actual outputs. The model links the demand for each journal type by library type in Module 2¹⁸ with an actual subscription price for each journal¹⁹ (calculated with reference to the sources of funds assumptions; and the proportion of total circulation / subscriptions accounted for by UK libraries).

Figure 4.2. Cost allocation by funding source and geographical location



* Non-academic subscriptions including corporate and individual users.

¹⁶ Like for other modelling estimation in the model this will be an assumption made for the total cost of journal production and by journal type, reflecting information (to the extent that we can access it) on the allocation of subscriptions by country¹⁶

¹⁷ Source: SCONUL

¹⁸ Assumes that list prices are available.

¹⁹ This also serves as a cross check for comparison of total cost per ‘notional’ journals with actual cost figures.

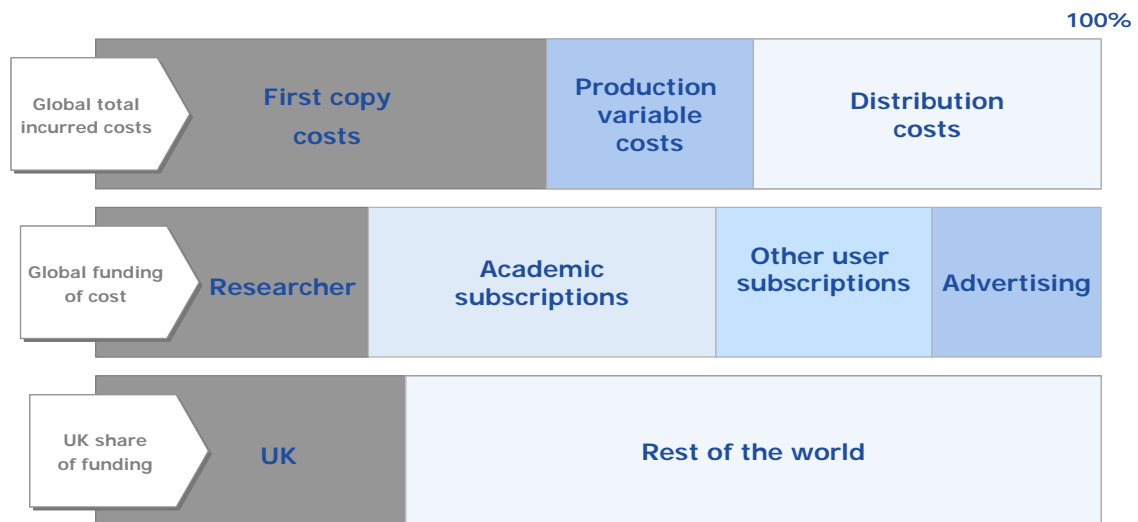
** Author-side payment includes contributions to open access publishing and author subsidies for publishing in form of page charges.

4.3.4. Module 3 Outputs

The *final* outputs generated by module 3 are shown in Figure 4.3²⁰ below. They include:

- The proportion of cash and ‘non-costs’ funded by different sources in the value chain. (This is the second ‘horizontal’ bar in Figure 4.3. The first bar shows the incurred cost allocation generated in module 1.)
- The proportion of funding of ‘non-cash’ and cash costs associated with the publication and distribution by the UK.
- A ‘top-down’ and ‘bottom-up’ estimate of the actual costs incurred by the UK academic libraries and academic research funders. (This is an allocation of the UK element of subscription funding, but is not shown explicitly in Figure 4.3)

Figure 4.3. Total journal publishing/distribution cost and funding allocation



4.4. Model calibration

As has already been noted, the model was populated with best estimate input assumptions and was calibrated against actual data to create a model **base case**. The main input assumptions were cross-checked to be largely in line with market data, for example: (i) total revenues generated by publishers; (ii) total actual UK (academic) library expenditure; (iii) actual distribution of journals in terms of circulation.

²⁰ Note that the Figure only refers to publication and distribution, since the access component is UK by definition. However, as indicated in Figure 4.2, these access costs need to be added to the publication and distribution costs to get the total UK cost of the system.

4.5. Scenario testing

Scenario testing can be carried out by changing input assumptions in the relevant input components of modules 1, 2 and 3. Although the model offers a scenario analysis and presents the main outputs in graphs showing variations to the base case, it is advisable to save different versions of the model to capture and compare different scenarios – this would also be the means by which modelling assumptions defining each scenario would be captured.

The model is designed to explore scenarios on most of the key input assumptions:

- **Unit cost scenarios** will allow the RIN to test the effects on total, average and marginal cost of changes in the fixed and variable cost per article. Unit cost scenarios will involve changing the assumed unit cost per article (fixed or variable) of specific activities.
- **Volume scenarios** will allow the RIN to test the effects on total, average and marginal cost of different volume of articles (and journals) flow through the system. Volume scenario testing will involve changing individual counts (allocations) articles between business models, journal types, subject areas, delivery models, publisher types, etc.

APPENDIX 1: BUSINESS MODEL VALUE CHAINS

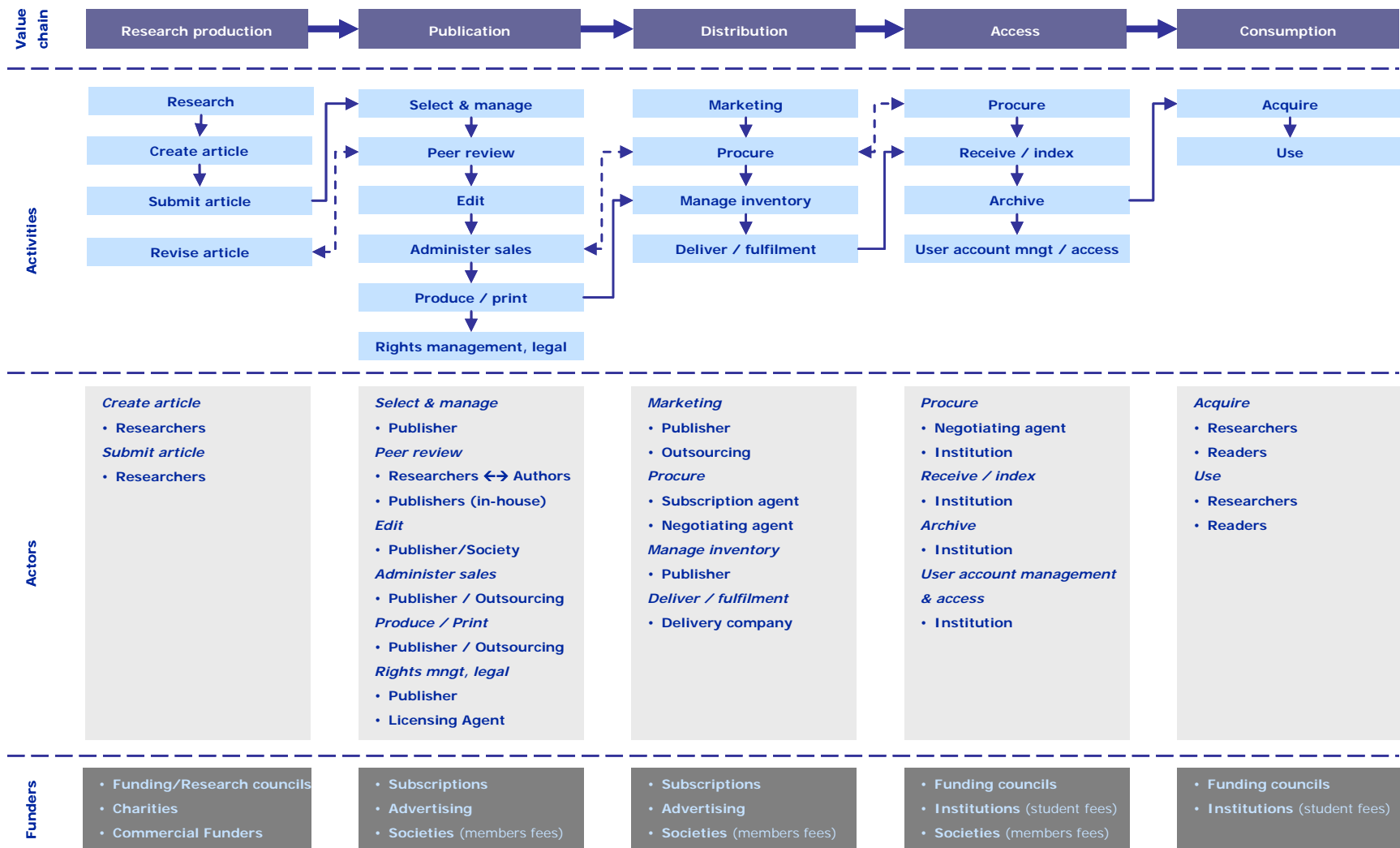
A1.1 Introduction

Our understanding of the value chain that underpins each business model is set out in the following diagrams. We considered the following basic business models.

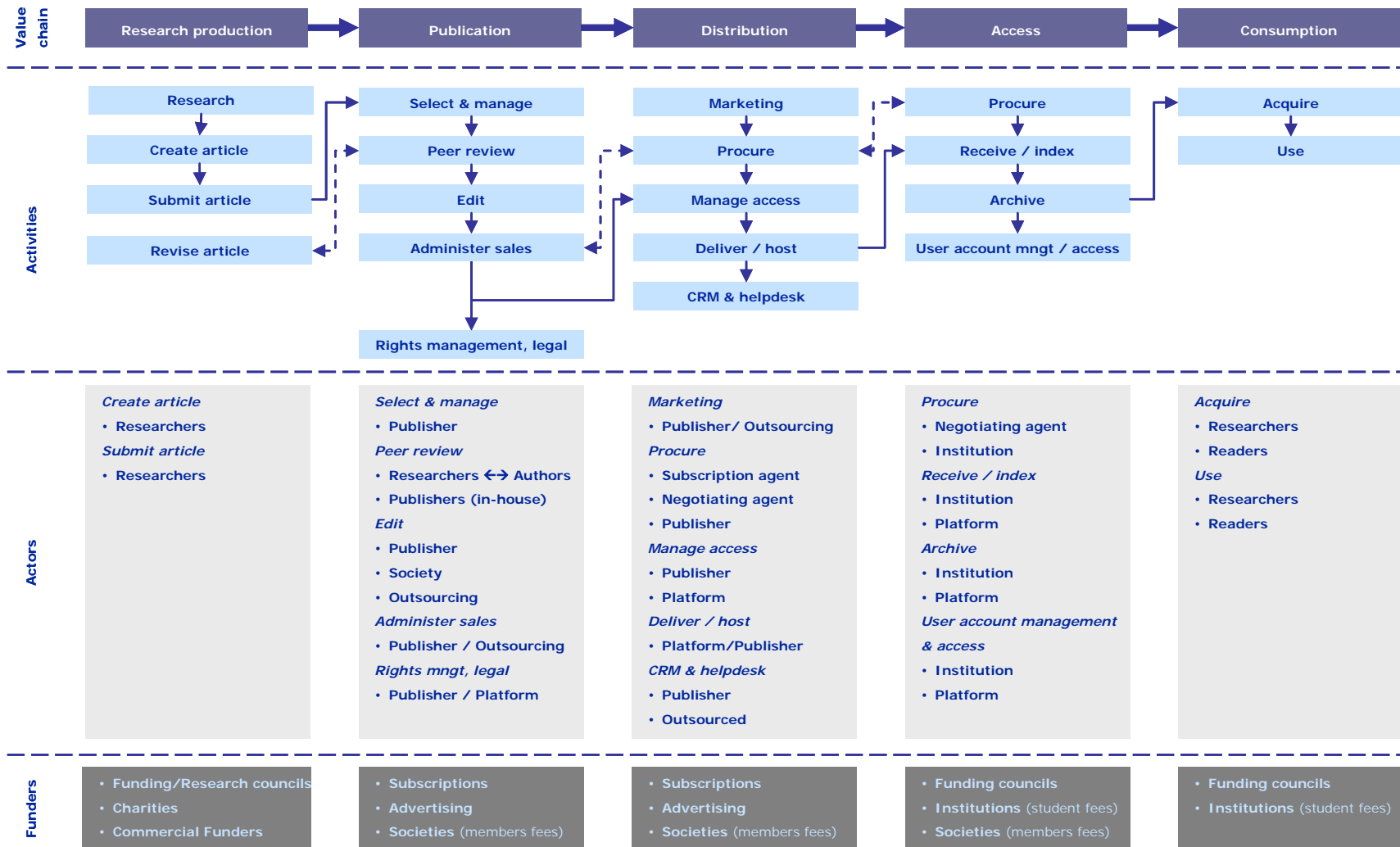
- subscription-based print journals;
- subscription-based electronic journals;
- open access journals; and

Each of the diagrams below sets out the activities that comprise the stages of the value chain, the key market participants / ‘actors’ and the principal sources of funding. The disaggregation of activities at each stage in the value chain formed an important basis for our approach to building up costs in the modelling work.

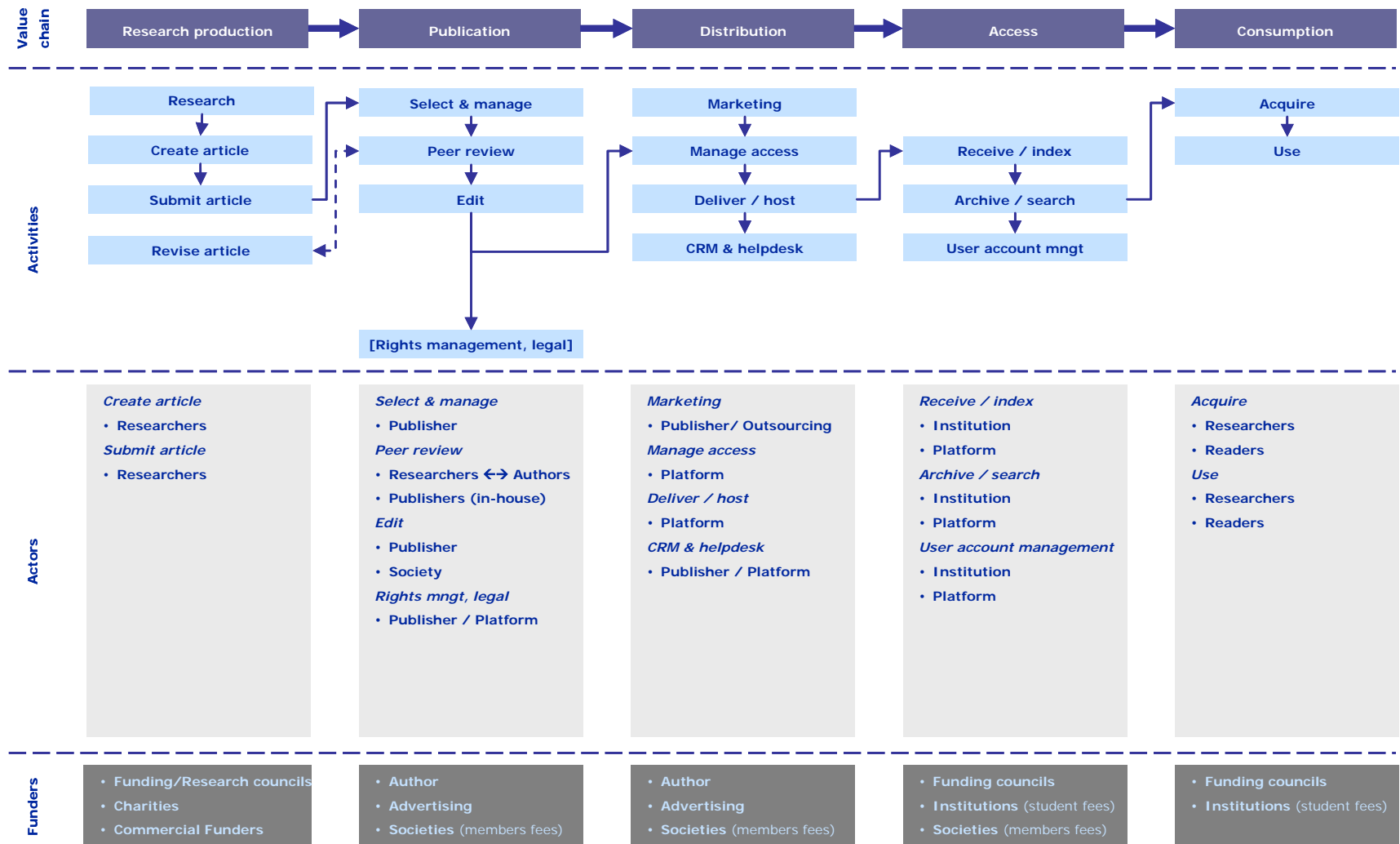
A1.2 Subscription-based print journal model



A1.3 Subscription-based electronic journal model



A1.4 Open access journal model



APPENDIX 2: COST DEFINITIONS

The categories of costs we define below broadly follow the definitions set out by King (2007). Unit costs will combine with flows of articles and flows of journals in the model to generate aggregated costs. Articles and journals are the main ‘volume drivers’ generating aggregated costs.

A2.1 Fixed & variable costs

In economics, **fixed costs** are defined as costs that do not change as output increases or decreases. In the context of the scholarly communications process, we consider there to be *per article* fixed costs, and *per journal* fixed costs – i.e. costs which are incurred regardless of the number of times an article or a journal is recreated (circulation). Per article fixed costs are most likely to relate to the publication component of the value chain, namely the cost of receiving each article, editing each article and (usually non-cash) peer-review costs.

Further *per journal* fixed costs are incurred in producing the journal, such as cover editing and journal marketing. Like per article fixed costs, per journal fixed costs will likely relate to publication – some authors refer to this as first-copy costs. Unlike print journals, much of this *per electronic* journal fixed cost is in distribution – the cost of electronic storage hardware.

We also consider there to be *per journal* **variable costs**. Variable costs change directly as output changes – in publishing per journal variable costs vary according to the circulation of the journal. Journals with a high circulation will have high total variable costs.

Variable costs are incurred with each additional copy of a journal produced. For a print journal they include print and distribution cost (e.g. paper, packing, postage) – per journal variable costs are lower, but high circulation will imply higher server capacity costs.

A2.2 Direct & indirect costs

We describe costs relating directly to the publication and distribution of, and access to articles or journals as **direct** costs. Examples of direct costs include the non-cash cost of peer reviewing an article, or the cost of hosting a journal using a third party web-hosting service. Direct costs exclude marketing, finance, HR and other costs not directly relating to the physical production of the article or journal itself.

It is useful to contrast direct costs with **indirect** costs. Indirect costs are the fixed, often company wide, cost of overheads, which are not attributable directly to the production of either an individual article or journal. Examples of indirect cost include management and administration cost, or depreciation of assets.

A2.3 Average cost (break-even price)

A journal publisher will break-even if he or she sets the price equal to the average cost of producing the journal. The average cost of producing journal type A is defined algebraically as:

$$\text{Average cost}_A = \text{Total cost}_A / \text{Circulation}_A \quad (1)$$

where,

$$\text{Total cost}_A = \text{Total fixed cost}_A + \text{Total variable cost}_A \quad (2)$$

where,

$$\text{Total variable cost}_A = \text{Variable cost}_A * \text{Circulation}_A \quad (3)$$

therefore, dividing (2) through by Circulation_A implies:

$$\text{Average cost}_A = \text{average fixed cost}_A + \text{variable cost}_A \quad (4)$$

Therefore the break-even price of a journal is equal to the average fixed cost plus the variable unit cost.