

Study on the economic and technical evolution of the scientific publication markets in Europe

Final Report - January 2006

Commissioned by DG-Research, European Commission

Undertaken by¹

Mathias Dewatripont, Victor Ginsburgh, Patrick Legros and Alexis Walckiers
ECARES, Université libre de Bruxelles

Jean-Pierre Devroey, Marianne Dujardin and Françoise Vandooren
Library Department, Université libre de Bruxelles

Pierre Dubois, Jérôme Foncel and Marc Ivaldi
IDEI, Université des Sciences Sociales, Toulouse

Marie-Dominique Heusse
Library Department, Université des Sciences Sociales, Toulouse

¹ The study has benefited from input by Bernard Balg, Laure Bessoles, Sabrina Bringuier, Elodie Camberlain, Ghislaine Chartron, Olivier Ertzscheid, Matias Labé, François Magnan, Marylène Poelaert and Jean-Michel Salaün

Table

of contents

EXECUTIVE SUMMARY	5
BACKGROUND AND GOAL OF THE REPORT	5
ECONOMIC AND POLICY CONTEXT	5
RESULTS OF THE ANALYSIS	6
POLICY RECOMMENDATIONS	11
1. INTRODUCTION	15
1.1. BACKGROUND OF THE STUDY	15
1.2. GOAL OF THE STUDY	154
1.3. SETTING THE SCENE: OVERVIEW OF THE ECONOMIC AND POLICY CONTEXT	16
1.3.1. <i>The economic and technological context</i>	16
1.3.2. <i>The policy context</i>	17
2. OVERVIEW OF THE MARKET FOR SCIENTIFIC PUBLISHING	21
2.1. MARKET IMPERFECTIONS	21
2.2. MARKET TRENDS UNTIL THE MID-1990'S	22
2.3. TRENDS SINCE THE MID-1990'S	24
3. ENTRY, PRICES, COSTS AND QUALITY	27
3.1. DESCRIPTION OF THE DATABASE	28
3.2. DESCRIPTIVE STATISTICS	32
3.3. PRICES AND QUALITY: AN ECONOMETRIC ANALYSIS	33
3.3.1. <i>A First, Simple Analysis of Prices, Citations and Ages</i>	33
3.3.2. <i>On Characteristics and Prices</i>	38
3.3.3. <i>Conclusions</i>	40
4. SCIENTIFIC PUBLISHING IN THE ELECTRONIC ERA: INSIGHTS FROM INDUSTRIAL ECONOMICS	43
4.1. TWO-SIDED MARKETS, NETWORK EFFECTS AND BUSINESS MODELS	43
4.2. NEW TECHNOLOGY AND COMPETITION	44
5. BUNDLING STRATEGIES	47
5.1. ECONOMIC LITERATURE ON BUNDLING AND DISCRIMINATION	47
5.2. BIG DEALS	48
5.2.1. Consortia	52
5.2.2. Lock-in Effects	52

6. VAT ISSUES	55
6.1. CURRENT SITUATION IN EUROPE	55
6.2. POSSIBLE SOLUTIONS	57
7. ACCESS TO RESEARCH OUTPUTS	59
7.1. ACCESS ISSUES	59
7.2. DEVELOPMENT OF OPEN ACCESS ARCHIVES	60
7.2.1. <i>Main advantages of open access archives</i>	61
7.2.2. <i>Questions raised about open access archives</i>	61
7.2.3. <i>Comments on concerns expressed by publishers</i>	63
7.3. OPEN ACCESS JOURNALS	64
7.3.1. <i>Facts</i>	64
7.3.2. <i>Competition in the Reader/Library-Pay and Author-Pay Systems</i>	67
7.4. RESEARCH FUNDING BODIES' POLICIES CHANGING CONDITIONS OF ACCESS	69
7.5. CONCLUSIONS	70
8. ELECTRONIC JOURNALS: PERMANENT ACCESS AND PRESERVATION	75
8.1. ONLINE ACCESS TO JOURNAL BACK ISSUES	75
8.1.1. <i>Born-digital journals</i>	75
8.1.2. <i>Digitized journals</i>	77
8.2. LONG-TERM PRESERVATION OF ELECTRONIC JOURNALS	78
8.2.1. <i>Technical developments</i>	79
8.2.2. <i>Roles and responsibilities</i>	79
8.3. CONCLUSIONS	81
9. STANDARDS AND INTEROPERABILITY	83
10. RECOMMENDATIONS	87
11. REFERENCES	91
11.1. ECONOMIC THEORY	91
11.2. SCIENTIFIC PUBLISHING	92
12. LIST OF ACRONYMS	97
ANNEX TO THE STUDY ON THE ECONOMIC AND TECHNICAL EVOLUTION OF THE SCIENTIFIC PUBLICATION MARKETS IN EUROPE	99

Table

of contents

EXECUTIVE SUMMARY

BACKGROUND AND GOAL OF THE REPORT

Dissemination and access to research results is a pillar in the development of the European Research Area. Aware of current public debates that reveal worries about the conditions of access and dissemination of scientific publications, the European Commission's Directorate-General for Research has commissioned a study that seeks: (i) to assess the evolution of the market for scientific publishing; and (ii) to discuss the potential desirability of European-level measures to help improve the conditions governing access to and the exchange, dissemination and archiving of scientific publications (taking into account all actors/stakeholders of the sector).

The report builds on a voluminous existing literature. It therefore updates the "state of the art" in terms of reports, studies, surveys, and articles. It has also benefited from considerable interaction with various actors/stakeholders of the sector, policy bodies, corporate associations and interest groups. Discussion meetings took place as well as participation and exchange in scientific conferences and policy forums. Three 'consultation days' were also organized, where preliminary results were discussed with publisher representatives, scholarly societies, research-funding organizations, and library representatives.

Scientific journals fulfill a double role of *certification* and *dissemination* of knowledge, the latter one including the archiving of knowledge to guaranteeing perennial access. It is the goal of this study to evaluate the ways in which these missions are fulfilled and whether in particular the public funds used to contribute to these missions are used in a cost-effective manner.

The report considers the specificities of the market for current journal issues. In doing so, it discusses the broad facts about the market; it undertakes a quantitative analysis of journal prices; it discusses the implications of technological innovation on pricing strategies

and the dynamics of entry; and it analyzes the implication of these developments in terms of competition policy.

It also discusses the various alternatives for disseminating and accessing scientific publications. This includes the question of access to research results on individual web pages or in public repositories, the development of open-access journals as well as other alternatives, such as pay-per-view, the question of the long-term preservation of electronic publications and the use of standards to ensure interoperability between systems.

The attention of public decision makers is required for two reasons. First it is well-established that science has a key role in fostering economic growth, and because scientific journals are an essential means of disseminating new knowledge in the academic community but also beyond. Secondly, much of scientific activity is publicly funded: the output of research is typically not bought by journals but 'donated' by publicly-funded researchers; so are to a large extent refereeing services for the evaluation of research; and finally, journals are bought by publicly-funded researchers or, more often now, by publicly-funded libraries. It is therefore crucial for public authorities to form a view on the relative efficiency of the scientific publication process.

ECONOMIC AND POLICY CONTEXT

The core STM (science, technology and medicine) publishing market is estimated between USD 7 billion and USD 11 billion, while in 2001 OECD countries allocated USD 638 billion to R&D.

In the last 30 years, the prices of scientific journals have been steadily increasing. Between 1975 and 1995, they increased 200%-300% beyond inflation. This was accompanied by a fall in subscriptions both by individual researchers and by libraries whose budgets got squeezed. Indeed, journal prices far outpaced the evolution of library budgets, which did increase at a somewhat slower pace than total academic research budgets.

As of 1995, publishers started to adopt digital delivery and provide online access to their journals. New technologies and the internet have dramatically improved the accessibility of scientific publications for researchers, but actual access to the literature still relies on their library's ability to pay the subscriptions. Digital delivery brought significant changes in the journal pricing policies (e.g. selling access to bundles of journals in multi-year contracts) and enabled the introduction of new business models (e.g. provide free access to journals by charging a publication fee). Libraries gathered into consortia to share the benefits of access and improve their bargaining positions in front of publishers. Technologies also enabled the development of alternative ways to disseminate and exchange scientific information more freely, such as open access e-print archives (i.e. server collecting papers deposited by their authors).

In view of the libraries' ongoing budgetary difficulties and of the opportunities provided by information technologies, and acknowledging the significant part of public funds involved in the scientific publishing process, a movement in favor of open access to scientific information has gained scale in the research community and research-related organizations. Declarations in favor of open access, such as *Budapest Open Access Initiative* and the *Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities*, were signed by thousands of individuals and by major research institutions and research funding bodies around the world. These concerns about access to research results have been echoed by the civil society (e.g. at the World Summit on the Information Society) and by political bodies at national and international levels (e.g. the *OECD Declaration on Access to Research Data from Public Funding*). In the UK, the House of Commons (2004) recommended that public funding agencies require open access to publicly-funded research through deposit of the publications in the authors' institutional repositories.

Following these declarations and recommendations, several important research funding bodies have established policies urging their funded researchers to publish in open access journals, offering to pay the publication fees if any, and/or to deposit their articles in an open access repository (e.g. the US National Institutes of Health, the UK Research Councils, the Wellcome Trust, CERN...). These policies have raised publishers and learned societies' concerns about the potential threat on their existence and activities: they fear that as articles become freely available in open archives and as search, access and retrieval facilities are enhanced by search engines and interoperability, journal subscriptions will be cancelled, thereby undermining the viability of their journals.

Starting from this global economic and research-policy context, this report provides an independent analysis of the conditions regarding access and dissemination of research results, with a view to maximizing societal returns on R&D investments.

RESULTS OF THE ANALYSIS

The report is subdivided in 8 sections, numbered 2 to 9. We follow this numbering here.

Section 2 starts by relying on the existing literature to stress the **broad facts about the market for journal publications**. It first discusses the significant difference between this market and the 'ideal perfectly competitive market'. Beyond the key role of public funding of authors, referees and journal purchases, it is worth stressing that this is an intermediated market, where libraries are the key buyers, which leads to lower reader price sensitivity. Moreover, it is a market where the best authors want to publish in highly-read journals and readers want to read journals which publish the best authors. This leads to 'virtuous circles' for journals, and to associated 'natural barriers to entry'.

An important question is to what extent these features allow publishers of high-quality journals to enjoy market power and set prices in excess of costs, a phenomenon that could be reinforced by ‘strategic barriers to entry’ if they are erected by incumbent publishers.

The broad market trends until the mid-nineties, and thereafter – where 1995 is the approximate start of the ‘electronic revolution’ are discussed next. Features that have been a constant from 1975 until today include:

- (i) the increasing reliance on journals as the main channel for dissemination of scientific knowledge, with a growth that parallels the growth of research produced;
- (ii) the dominance of the ‘reader-pay’ or ‘library pay’, as opposed to the ‘author-pay’ model of journal dissemination;
- (iii) the existence of many publishers in the market, with two big groups of publishers: For-profits (FP) and Not-for-profits (NFP), the latter group including learned societies and university presses;
- (iv) the very fast growth of some big FP publishers, through new journal introduction, through the running of journals from learned societies, and through mergers.

The technology of publishing has always implied high fixed costs. The transition to the electronic era has implied a fall in dissemination costs but a one-off (fixed) transition cost, as well as high fixed costs but low marginal costs of providing new services (citation data, cross references, ...). It has thus implied increased returns to scale but also lower entry costs for new journals. This has allowed some experimentation with new business models, like open-access, author-pay journals.

The market has witnessed a sharp increase in prices: In the 1975-95 period, FP journal prices have grown by *300% beyond inflation*. The rate has been almost the same for learned societies, although from much lower initial level. And for university presses, it has been *+200% beyond inflation*, from an even lower initial level.

These increases have been accompanied by a fall in subscriptions, especially by individual researchers but also multiple library subscriptions.

After 1995, prices have kept rising faster than inflation, but less so than before (given the limits of library budgets). Moreover, there have been big gains in access, and thus improvements in ‘value for money’ thanks to the ‘Internet revolution’. A key change that accompanied this revolution has been the move from individual journal pricing to bundled pricing. Bundled pricing is less transparent because it is very often based on the libraries’ pre-existing print subscriptions, and individual electronic journals are usually priced as a percentage of their print price (typically 90% of print subscription). In addition, many libraries continue subscribing to print copies along with the electronic versions because of VAT taxation differentials, discussed in Section 6. Another development has been the formation of ‘library consortia’ which bargain with big publishers on behalf of individual libraries. This led to so called ‘Big Deals’.

Section 3 presents a detailed **quantitative investigation of journal prices**. It concerns 22 fields of (natural and social) sciences with the largest members of journals, as recorded by the Journal of Citation Reports. In total, around 2,700 journals are considered. Four conclusions can be highlighted:

- (i) FP publishers have been the main impetus behind the growth of publishing activity. NFP publishers – and especially learned societies – have focused on higher quality journals (where quality is defined as the number of citations to the journal divided by the average number of citations to a journal in the same field).
- (ii) Prices per journal or per article vary quite considerably between scientific disciplines.

- (iii) Prices per journal or per article, are significantly higher for FP than NFP journals. Specifically, for a given scientific field, a given level of quality (as measured by citation counts), and a given age of the journal, an FP journal (or FP journal article) is on average three times as expensive as an NFP journal (or NFP journal article).
- (iv) There is a statistically significant positive link between the quality of a journal and its price, especially for FP journals.

Fact (i) indicates that FP publishers have been important in accompanying the growth of research. Facts (ii) and (iii) imply that substitution possibilities across journals are limited, so that publishers do have significant market power. Fact (iii) is consistent with prior evidence, but is documented more systematically in this report.

Fact (iv) sheds light on the pricing behavior of publishers. It has been argued that higher prices for FP publishers (which is on average three times the price charged by NFP publishers, for a given field, citation count and age of the journal) could be the consequence of higher costs, due to increasing returns to scale in publishing combined with lower circulation levels for FP journals, especially in comparison with learned-society journals. While it is true that average circulation is lower for FP journals, the positive link between prices and citation count is not consistent with cost-based pricing. This report argues that higher citation counts should be associated with lower average publishing costs, because they should be positively linked with higher circulation. The link is instead consistent with ‘value-based pricing’, again an indication of market power.

Note that the analysis is based on individual journal prices, that is, on prices for individual journals in printed form, which are publicly available, while ‘bundled prices’ that are now common are not public. Since individual print prices remain the basis for bundled pricing, they are nonetheless informative. It is however important to investigate the consequences of this move towards bundled pricing.

Given the lack of publicly available data, the report relies on theoretical insights from industrial economics. Further analysis, based on information about actual bundled prices, would obviously be welcome.

Section 4 discusses the **impact of the ‘electronic revolution’ on the certification and dissemination of research.** It emphasizes the potential for alternatives to existing scientific journals in performing research dissemination, through individual web pages or repositories (helped by search engines) or through ‘open-access journals’. It also stresses the strong position of publishers who control access to a sizable research portfolio that is considered vital by researchers in order to perform high-quality research.

As emphasized in **Section 5**, such a position is further strengthened by ‘Big Deal’ pricing strategies made possible by **bundling large portfolios of journals.** Bundling has emerged at the same time as electronic publishing, and also at the same time as the formation of ‘library consortia’, whereby libraries of a given geographical area (and/or subject) join forces in order to bargain with publishers. This ‘buyer concentration’ remains however modest in comparison with publisher concentration: the largest library consortium represents 2 or 3% of global journal purchases, while the largest publisher represents more than 20% of journal sales. Since researchers do not see the various publishers as good substitutes and need access to all good journals, consortia only introduce a relatively weak ‘buyer-power’ counterpart to the rising concentration in the publishing market.

Consortia include members that are heterogeneous (large and small universities, more or less research-oriented universities). ‘Big Deal’ practices are beneficial to consortium members to the extent that they allow sharing resources and providing more flexible access to journals. But they also allow publishers to extract more revenue from consortium members, by making it expensive to cancel preexisting subscriptions.

Since these contracts are moreover signed for significant periods of time, they tend to rigidify library budgets and make it difficult for potential entrants to profitably enter the market. This effect of bundling as a deterrent to entry turns out to be the key problem identified by the industrial economics literature. Entry deterrence reduces the desirable diffusion of knowledge and is particularly worrisome at a time at which the electronic revolution makes it especially desirable to leave room open for new entry and new business models.

This report therefore stresses the importance of promoting pricing policies that limit barriers to entry. Specifically, (i) there should be no penalties for net journal cancellations; (ii) prices of electronic access (within or without consortia) should not depend on the historical number of print subscriptions, but should instead be related to actual usage (or number of faculty and students) by institutions; (iii) increased usage over time should not lead to price increases if publisher costs do not increase as a result of the rise in usage.

Section 6 considers the bias introduced by the higher **VAT rate** that applies to electronic journals vs. their printed version, and lead libraries to subscribe to both. This section describes the current situation and underlines that the higher rate applied to electronic delivery of information in Europe strongly affects European research institutions, especially when compared to other countries where electronic services are exempt from tax, such as in the USA. It examines two possible solutions: (i) either apply a reduced VAT rate on electronic scientific information, which requires an amendment of Annex H of the VAT Directive; (ii) or introduce a tax refund mechanism for research institutions, as in Sweden or Denmark.

Section 7 deals with the important issue of the **timing of access to current research results** and discusses the tradeoff between the advantage of immediate and efficient access and its consequences that would prevent publishers from legitimately charging readers for their output and services, and eventually lead to the disappearance of the reader/library-pay model.

The evidence discussed in Sections 3 to 5 points however to a market where publishers do enjoy market power. This means that access can be improved significantly before one has to start worrying about threatening this reader/library-pay model. At present, immediate access can be achieved through individual web pages, through open access repositories or archives collecting e-prints deposited by their authors, and through open access journals which rely on various sources of income. In some countries, public authorities have funded large-scale projects to develop portals providing free online access to selected scholarly journals published in their countries (e.g. SciELO in Latin American countries, J-STAGE in Japan, various projects in India). As an increasing volume of research output from outside Europe becomes openly accessible, it raises the question of the visibility and accessibility, and of the subsequent potential impact, of European research as most articles by European researchers are published in subscription-based journals. In this respect, policies like those put in place by the U.S. NIH, the UK Research Councils or the Wellcome Trust – which promote *open availability of research results no later than 12 months after publication* – could be emulated across Europe, after discussions with the publishers. The aim could be to ensure that published European-funded research (at EU, national, or regional levels) be deposited in standardized open archives some time after publication. Recent surveys show that a majority of researchers seem willing to self-archive their articles if induced to do so by their employer or funding body.

Section 7 also discusses more broadly the **future of the scientific publishing industry**. While the reader/library-pay model is still predominant at the moment, it discusses existing author-pay journals and stresses that experimentation is taking place. It points to the possibility of a third model, where the reader would ‘really’ pay, that is, would be charged per article download. All three models have costs and benefits, and at this point it is too early to draw conclusions on desirable evolutions (whether it involves the dominance

of a single business model or the coexistence of several models). In comparison with the current reader/library-pay model, both the author-pay and the pay-per-download models would raise price sensitivity – this is especially true of the author-pay model, since substitution possibilities among journals are higher for authors than for readers – and could therefore be expected to lower prices and raise access to knowledge. On the other hand, monitoring downloads would entail a cost, and there are concerns about whether reputational incentives will be sufficient to ensure quality maintenance in author-pay journals (i.e. to prevent the lowering of scientific standards as a way to raise short-term revenue by publishing more papers).

The industry is fast changing and from discussions with the actors it seems clear that the future is still uncertain. Most of the economic analysis has been within the traditional distribution system, based on subscription and negotiations between publishers and libraries. The prevalence of this system makes it difficult to evaluate the feasibility of alternative systems, principally because such feasibility depends on how the overall budget for publication and diffusion is allocated. It is worth noting that, if the research funding authorities want to ‘give a chance’ to the author-pay model, they have to allow for a ‘level-playing field’ in comparison with the reader/library-pay model, that is, provide funding for publication costs and not only for library budgets.

The electronic revolution has led not only to increased access but also to the increased availability of statistics, such as journal rankings based on citation counts and impact factors used in evaluating researchers. It would be natural to complement these by *more comprehensive journal rankings* that would take explicitly into account other dimensions, such as speed of publication, self-archiving allowed upon publication or within a short delay, copyright policies, abstract and indexing services, reference linking, long-term preservation provisions etc.

Citations and impact factors should however remain the predominant dimension of journal quality for scientific evaluation.

From the discussion in this section, it appears that there is a central role for funding bodies to define policies which will improve access and dissemination of publications, especially in terms of self-archiving requirement, copyright provisions and more comprehensive rankings of journals accounting for dissemination quality.

Section 8 discusses issues related to access to the ‘stock of knowledge’. It first stresses that, in the electronic era, publishers, instead of selling new issues of the printed journal, rent electronic access to the entire journal collection. Conditions of **continued access to electronic journals** that were previously subscribed vary from publisher to publisher and can entail a significant cost, thereby strengthening publisher market power. Furthermore, digitisation of printed journals is costly and the availability of journals digital archives varies according to the size and type of publishers, domain, language, and national policies: while publishers with large portfolios tend to digitise their complete collections and license access to their archives, smaller publishers are more likely to subcontract with non-profit organisations. Publicly-funded initiatives and public-private partnerships are also emerging in non-English speaking countries to digitise quality journals in Social Sciences and Humanities in order to promote their national research output.

Beyond this, **long-term preservation of electronic publications** is a problematic issue, in a world where publishers may come and go. Such long-term preservation is a public-interest mission that may require public-sector commitment and funding, since one cannot expect or wish private actors to earn a sufficient return on this activity. JSTOR, an independent not-for-profit organization that provides electronic access to the backfiles of scientific journals could be a natural model to start from for a European initiative in this respect.

Finally, **Section 9** considers **standards** that could ensure **interoperability** between different hardware and software systems, making it possible for data to be transferred between proprietary applications, to be searched and viewed by users through the internet, and to be linked across platforms. It highlights some of the standards and technologies currently used for electronic publications and underlines some future developments that could be supported by research funding programmes. It points to synergies with current projects-supporting programmes (European Directorate-General Information Society, JISC in the UK), with the forthcoming European "i-2010: Digital Libraries" Communication on scientific information and with the overall e-infrastructure building strategy for the European Research Area.

POLICY RECOMMENDATIONS

Section 3 has documented the very significant price differences – from 1 to 3 between non-profit and for-profit publishers for comparable-quality journals (or articles). This dispersion is indicative of market power, as is the positive link between prices and number of citations. Since citations should be positively correlated with circulation, and thus with lower average costs (given the scale economies, due to high ‘first-copy costs’), this positive link is a sign of ‘value-based pricing’ rather than ‘cost-based pricing’, and thus again of market power.

While it is important to stress the societal value of the existing publication system, it is also important to acknowledge the societal cost linked to high journal prices, in financial terms for public budgets, but also in terms of limits on the dissemination of knowledge and therefore of further scientific progress. This can be particularly unfortunate at a time where technological possibilities due to the Internet have the potential to dramatically raise dissemination possibilities.

Consequently, policies should make sure that the market is sufficiently competitive and ‘dissemination-friendly’. In particular, they should address the need to:

- (i) enhance access to research output;
- (ii) prevent strategic barriers to entry and to experimentation.

The first five recommendations concentrate on access issues, which are currently the subject of significant policy discussions.

RECOMMENDATION A1. GUARANTEE PUBLIC ACCESS TO PUBLICLY-FUNDED RESEARCH RESULTS SHORTLY AFTER PUBLICATION

Research funding agencies have a central role in determining researchers’ publishing practices. Following the lead of the NIH and other institutions, they should promote and support the archiving of publications in open repositories, after a (possibly domain-specific) time period to be discussed with publishers. This archiving could become a condition for funding.

The following actions could be taken at the European level: (i) Establish a European policy mandating published articles arising from EC-funded research to be available after a given time period in open access archives, and (ii) Explore with Member States and with European research and academic associations whether and how such policies and open repositories could be implemented.

RECOMMENDATION A2. AIM AT A ‘LEVEL-PLAYING FIELD’ IN TERMS OF BUSINESS MODELS IN PUBLISHING

There is a central role for education and research funding authorities in the shaping of new models for publishing and communicating research results. They should be aware that the rules governing education and research budgets have strong implications for the viability of various business models. At this point, it seems desirable to allow for experimentation and competition between various possible business models, which means allocating money to libraries to subscribe to reader or library-pay journals but also to authors to pay for publication costs in author-pay journals, and to researchers in the reader-pay model. Establishing relative priorities in this respect should become a key policy debate.

RECOMMENDATION A3. 'EXTENDED QUALITY' RANKINGS OF SCIENTIFIC JOURNALS

This recommendation aims at raising researcher awareness of journal quality beyond scientific quality, *stricto sensu*. While scientific quality, approximated for example by citation counts, should remain the dominant criterion, dimensions related to the **quality of dissemination** (self-archiving authorisation, publisher archiving provisions, copyright provisions, abstracting and indexing services, reference linking, etc.) could be tracked explicitly and possibly valued by research funding bodies. There could be an impetus from public authorities at the European level for such an initiative, which would naturally induce publishers to stress good practices in these dimensions.

RECOMMENDATION A4. GUARANTEE PERENNIAL ACCESS TO SCHOLARLY JOURNAL DIGITAL ARCHIVES

Given the heterogeneity of the publishers' current provisions, promote the creation of not-for-profit long-term preservation archives, which balance interests among publishers, libraries, and scholars. More particularly (i) Promote business models for legal-deposit libraries to allow remote online access to their journals digital archives, therefore providing them with return on investments and making the preservation efforts cost-effective; (ii) Investigate the feasibility/desirability of the creation of a European non-profit journals preservation organisation ("JSTOR-like") and of other subject-based archives in relevant domains; (iii) Determine the standards under which archives must be accessible and set up a portal as a central access point to digital journals and articles.

RECOMMENDATION A5. FOSTER INTEROPERABLE TOOLS TO IMPROVE KNOWLEDGE VISIBILITY, ACCESSIBILITY AND DISSEMINATION

This could be achieved by (i) supporting research and development on interoperability issues, notably on metadata to improve scientific information search and retrieval efficiency and

on the XML format to improve and accelerate the overall publishing process, and by (ii) promoting the wide implementation of linking technologies, especially the open standard OpenURL, and of interoperable standard protocols, especially the OAI-PMH that enables metadata harvesting and searching across different platforms. Both developments could be taken into account by the European Commission in its e-infrastructure building strategy for the European Research Area (involving DG-Information Society R&D funding programmes and the forthcoming "i2010: Digital Libraries" Communication on scientific information).

The next three recommendations concern the need to prevent strategic barriers to entry and to experimentation and also excessive concentration.

RECOMMENDATION B1. PROMOTE PRO-COMPETITIVE PRICING STRATEGIES

The key issue identified in terms of market access concerns pricing policies, and more particularly the lock-in effect associated to 'Big deals'. Specifically, the limited savings libraries obtain for net subscription cancellations does make it hard for newcomers to have access to library budgets. The following simple rules (that could be promoted by the European authorities) would avoid some of the long term negative effects of big deal contracts on entry and competition: (a) The price of the electronic access should not depend on the historical number of print subscriptions; (b) Prices should be related to transparent indicators, like usage or the number of faculty, students, etc., as is the case with JSTOR for instance; (c) Libraries should have the possibility to choose among variable dimension bundles, and compose their preferred bundle. Therefore, journals in a bundle should also be priced individually, and prices of bundles should ideally be made public; (d) Finally, note that overall usage has been on the rise thanks to the Internet, and can be expected to keep growing at least for some time. One should avoid having prices increase with such usage as long as publishing costs do not increase as a result of this rise in usage.

RECOMMENDATION B2. SCRUTINIZE FUTURE SIGNIFICANT MERGERS

The market has become more concentrated due in part to acquisitions by large for-profit publishers, and some of the price increases can be traced back to these mergers, though the largest firm controls less than 30% of the overall market (market shares are however higher in some scientific fields). It has been shown that publishers with large journal portfolios have an incentive to set higher prices. This indicates that further acquisitions by large publishers should be scrutinized by the relevant European authorities.

RECOMMENDATION B3. PROMOTE THE DEVELOPMENT OF ELECTRONIC PUBLICATIONS

Eliminate the unfavourable tax treatment of electronic publications, by (i) either applying a reduced VAT rate to all types of scientific information, whether print or electronic; (ii) or, given the political difficulty of implementing this solution which requires unanimity of Member States, by introducing a tax refund mechanism for research institutions, as is already the case in Sweden and Denmark. Though the European Commission could play a role in supporting and promoting either solution, this decision is ultimately left to the Member States.

Encourage public funding and public-private partnerships where there is little commercial investment in the creation of journals digital archives, especially for quality European journals in Social Science and Humanities. Such initiatives require further investigation of the structure and organisation of publishing markets for SSH journals which are quite different between countries, especially the distribution among private and public actors.

Finally, the last two recommendations stress the need for further discussions and study concerning this important market.

RECOMMENDATION C1. SETTING-UP AN ADVISORY COMMITTEE

Discussions with all the stakeholders during the study made it clear that regular contacts are necessary, since the industry's practices are moving very fast, and will keep doing so in the future. We advise to set up a committee composed of the various interested parties: publishers, librarians, funding bodies, authors and researchers, who should be responsible for observing practices, meeting (say once or twice a year) to discuss and recommend changes if need be, and reporting the results of the discussions to the Research (and possibly the Competition) DG's of the European Commission.

RECOMMENDATION C2. FURTHER INVESTIGATION

This study is obviously not exhaustive. Here are some topics where further investigation could be commissioned:

- A first important topic concerns the evolution of **copyright provisions**, which we address only briefly in this report. While publishers have become more permissive over time, in particular in terms of the posting of published material on individual web pages, it would be good to investigate precise legal solutions that would provide legal certainty to authors, but also potentially to other parties, in terms of dissemination of published material.
- A second topic concerns the **economic analysis of alternative forms of dissemination**: for example, the feasibility/desirability of alternative publishing business models (pay-per-download, author-pay systems, hybrid systems) and of the unbundling of certification and dissemination; and the long-term sustainability of open repositories.
- Finally, a third topic concerns **technological developments**: Research could be supported for example on interoperability issues and on the specifics of long-term preservation issues.

1. INTRODUCTION

1.1. BACKGROUND OF THE STUDY

Dissemination and access to research results is a pillar in the development of the European Research Area. Aware of current public debates that reveal worries about the conditions of access and dissemination of scientific publications, DG-Research has commissioned a study that seeks: (i) to assess the evolution of the market for scientific publishing; and (ii) to discuss the potential desirability of European-level measures to help improve the conditions governing access to and the exchange, dissemination and archiving of scientific publications (taking into account all actors/stakeholders of the sector).

The team that has undertaken the study is multifaceted, consisting in: (i) Economists from ECARES (the European Centre for Advanced Research in Economics and Statistics of the Université Libre de Bruxelles – ULB) and IDEI (the Institut d’Economie Industrielle of the Université des Sciences Sociales de Toulouse – UT1) and (ii) Information scientists from ULB and UT1 libraries. The team did also benefit from the help of associated academics from law and information sciences.

In terms of information resources and contacts, the study first builds on a voluminous existing literature. It therefore updates the “state of the art” in terms of reports, studies, surveys, articles... The study also presents data collected for the purpose of an economic analysis, in particular about journal prices and journal citations. The study has also benefited from considerable interaction with various actors/stakeholders of the sector, be they policy bodies, corporate associations and interest groups. Discussion meetings took place as well as participation and exchange in scientific conferences and policy forums. Three ‘consultation days’ were also organized, where preliminary results of the study were discussed respectively with: (i) publisher representatives; (ii) scholarly societies and research-funding organizations; and (iii) library representatives.

The study focuses on the “*Scientific publication markets in Europe*”. The bulk of it concerns English-language STM (‘Science, Technology and Medicine’) journals, which is the primary focus of the study, even if some specificities of Social Sciences and Humanities journals and of journals published in other languages than English are also briefly mentioned.

The study first considers the specificities of the market for current journal issues. In doing so, it discusses the broad facts about the market; it undertakes a quantitative analysis of journal prices; it discusses the implications of technological innovation on pricing strategies and the dynamics of entry; and it analyzes the implication of these developments in terms of competition policy.

The second part of the study is entirely devoted to the various alternatives for disseminating and accessing scientific publications. This includes the question of access to research results on individual web pages or in public repositories, the development of open-access journals, and the question of the long-term preservation of electronic publications. Some of these discussions have implications in terms of Intellectual Property Rights, which are only briefly discussed in the context of this study.

1.2. GOAL OF THE STUDY

This study evaluates the evolution of the market for scientific publishing, and in particular on its central component, that is, scientific journals. The stakes of such an evaluation are significant, because it is well-established that science has a key role in fostering economic growth (see among others Aghion and Howitt, 1998), and because scientific journals are an essential means of disseminating new knowledge in the academic community but also beyond (as documented by Stephan, 1996, and Tenopir and King, 2000).

The scientific publishing process impacts very much the two pillars of scientific progress, that is, the *creation* and the *dissemination* of new

knowledge. Indeed, researchers want to publish their own research in scientific journals not only to have their findings widely *distributed* but also to have them *certified* in terms of quality, which helps both attracting the attention of fellow researchers and also demonstrating the value of the research, which in turn helps the researcher getting promotions and/or outside job offers. Scientific journals thus fulfill a double role of *certification* and *dissemination* of knowledge, the latter one including the archiving of knowledge to guaranteeing perennial access. It is the goal of this study to evaluate the ways in which these missions are fulfilled and whether in particular the public funds used to contribute to these missions are used in a cost-effective manner.

The stakes for public-decision makers are very significant indeed: Not only is the output of the scientific publishing process of crucial importance for society's development, but its cost is also very relevant, because so much of its activity is publicly funded: the output of research is typically not bought by journals but 'donated' by publicly-funded researchers; to a large extent, refereeing services for the evaluation of research outputs are also heavily subsidized by publicly-funded academic institutions; and finally, journals are bought by publicly-funded researchers or, more often now, by publicly-funded libraries. It is therefore crucial for public authorities to form a view on the relative efficiency of the scientific publication process.

Thus, we take the goal of this study to be to provide an evaluation of the scientific publication market in terms of its provision of incentives to adequately create and disseminate scientific knowledge. Concretely, we are therefore interested in the ability of this market, including the alternative models enabled by the new technologies, to offer cost-effective and high-quality levels of certification, dissemination and archiving services to the research community.

1.3. SETTING THE SCENE: OVERVIEW OF THE ECONOMIC AND POLICY CONTEXT

1.3.1. THE ECONOMIC AND TECHNOLOGICAL CONTEXT

As publication has a key role in the overall research process, the number of journal articles published has been roughly proportional to the growth in the number of scientists. Unsurprisingly, scientific publishing has also become a significant economic activity. The core STM publishing market is estimated between USD 7 billion and USD 11 billion (OECD 2005; note that in 2001 OECD countries allocated USD 638 billion to R&D).

In the last 30 years, a constant topic of debate has been the price of scientific journals, which has been steadily increasing. In the period 1975-1995, dubbed the "serials crisis", journal prices increased 200%-300% beyond inflation. This price evolution was accompanied by a fall in subscriptions by individual researchers and libraries. Indeed, journal prices far outpaced the evolution of library budgets, which in fact did not increase as fast as total academic research budgets. The pressure on library budgets resulted in reduced journal subscription as well as crowding out of some alternatives like monographs.

As of 1995, building upon ICT developments, publishers started to adopt digital delivery and provide online access to their journals through high-performance search platforms. New technologies and the internet have dramatically improved the accessibility of scientific publications for researchers, but actual access to the literature still relies on their library's ability to pay the subscriptions. Indeed, journal prices have kept increasing faster than inflation, though at a slower rate than during the previous 20 years, and electronic access fees bear an additional burden on libraries' budgets.

Statistical data on European libraries, from 1997 to 2001, shows a decrease in the numbers of books and periodicals acquired, while at the same time there has been a large increase (28%) in spending on acquisitions, including on electronic information (Fuegi and Jennings, 2004). Digital delivery enables the introduction of new business models and brings significant changes in the journal pricing policies: one has moved from individual pricing and selling of journals to “Big Deals”, i.e. selling bundles of journals whose prices vary from institution to institution; from annual subscriptions to licenses lasting several years. Libraries have gathered into consortia to share the benefits of access and improve their bargaining positions in front of publishers. As a short term result, the number of electronic journals licensed by libraries has been increasing over the last couple of years. But the legacy of the previous price increases is still there and many libraries face difficulties in renewing ever costlier licenses (the University of Cornell, for instance, renounced a Big Deal contract in 2004 in order to be able to cancel subscriptions to expensive journal titles).

In parallel to the publishing industry’s efforts to develop high-performance platforms for searching e-journals, scientists have continuously built upon ICT developments to improve communication and research collaboration processes (they have invented the Internet), by means of e-mail discussion lists, creating data bases, sharing models and computer programmes etc. In the nineties, the first e-print archives (server collecting papers deposited by their authors²) and free online peer-reviewed journals emerged, providing free access to research articles. Indeed, the technologies enable the development of alternative ways to disseminate and exchange scientific information freely and efficiently, as well as the creation of new business models to provide open access to journals, for instance by charging a publication fee (the “author-pay” model). Starting from isolated initiatives, the movement in favour of open access to scientific information has gained scale through the mobilization of groups of researchers and libraries, the development of open access article archives or repositories, and the recent policies of research funding bodies about access to research output.

1.3.2. THE POLICY CONTEXT

In view of the libraries’ difficulties and of the opportunities provided by information technologies, and acknowledging the significant part of public funds involved in the scientific publishing process, awareness of and concerns about scholarly communication issues have been rising in the research community and research-related organizations.

In 1998, SPARC, the *Scholarly Publishing and Academic Resources Coalition*, an alliance of universities, research libraries and organizations, started to raise awareness about scholarly communication issues and to support actions enhancing broad and cost-effective access to peer-reviewed scholarship. In 2001, the Public Library of Science, a coalition of researchers, launched a petition requesting that all published articles be freely available within six months of publication; it was signed by roughly 34000 scientists from 180 countries³. Soon after, three major public statements of international scope were issued in favour of “open access to scholarly journal literature”:

- the [Budapest Open Access Initiative](#)⁴ (February 2002), signed by more than 4000 individuals and organizations from around the world, representing researchers, universities, laboratories, libraries, foundations, journals, publishers, learned societies;
- the [Bethesda Statement on Open Access Publishing](#)⁵ (June 2003), issued by scientists, libraries and publishers agreeing on significant, concrete steps that all relevant parties can take to promote open access publishing.
- The [Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities](#)⁶ (October 2003), signed by more than 60 organisations including large national research institutions such as France's CNRS and Germany's Max-Planck Institutes; national Academies of Science such as those of China, India and the Netherlands; international research institutions such as CERN; and individual universities and research funding agencies around the world.

Although these three statements differ from one another in some ways, they all agree that the scientific literature (peer-reviewed articles or preprints) should be freely accessible to all users with an Internet connection, and there should be no restrictions on how published material is subsequently used except to require that proper attribution of the work be given to the author and that the authors retain control over the integrity of their work.

The [Wellcome Trust](#), the UK's biggest non-governmental funder of life sciences research, commissioned two reports in 2003-2004 on the economics of scientific research publishing, investigating costs and business models, in order to understand how it acts to influence the dissemination of the research it funds. The reports led the organization to definitely support the publishing of research results it funds in an open access format⁷.

The research community's concerns about access to research results have been echoed by the **civil society** and by **political bodies** at national and international levels:

- The Declaration of Principles of the [World Summit on the Information Society](#), (Geneva 12 December 2003) stated: "We strive to promote universal access with equal opportunities for all to scientific knowledge and the creation and dissemination of scientific and technical information, including open access initiatives for scientific publishing".(Article 28⁸). In a UNESCO thematic meeting held in May 2005, recommendations were issued for the practical implementation of the WSIS Open Access resolution, urging UNESCO and other UN specialized agencies to support open access repositories⁹.
- Governments expressed their concerns regarding access to research data. In January 2004, more than 30 nations signed the [OECD Declaration on Access to Research Data from Public Funding](#), recognising that "fostering broader, open access to and wide use of research data will enhance the quality and productivity of science systems worldwide"¹⁰.

Since then, the OECD has published a report that outlines a number of policies and initiatives which could enhance the digital delivery of scientific and technical information. It stresses that: "Access to public and government-funded research content is a crucial issue, and there is considerable potential for governments to provide a lead in enabling digital delivery and enhanced access to publicly funded scientific and technical information"¹¹.

- In the US, the "[Public Access to Science Act](#)" introduced in 2003 would require research substantially funded by the federal government to be made available in the public domain¹². In 2004, a large number of U.S. public-interest groups launched the [Alliance for Taxpayer Access](#)¹³ to support open access to taxpayer-funded research.
- In the UK, the [House of Commons Science and Technology Committee](#) issued a lengthy report in July 2004, recommending that public funding agencies require open access to publicly-funded research through deposit in the authors' institutional repositories¹⁴.
- In Finland, in its report issued in March 2005, the "Open Access Scientific Publishing Committee" appointed by the Ministry of Education recommended to research funding agencies and organisations conducting research that they set up open access archives and encourage researchers to deposit copies of their publications in these open archives¹⁵. Similar statements have been issued for example in Norway and Denmark¹⁶.

The "open access" debate has progressively broadened to involve all public actors concerned with research results dissemination and communication. There is growing concern that "publicly funded research data are a public good, produced in the public interest, and as such they should remain in the public realm"¹⁷. This declaration applies to all research results, data and literature, as scientific and technical advances are made possible only by the sharing of research results¹⁸. The Berlin Declaration extends the open access model to all original scientific results, including "raw data and metadata, source materials, digital representations of pictorial and graphical materials and scholarly multimedia material".

Even in the face of terrorism, the US National Research Council considered that the benefits of open access to genome data on pathogens for maintaining progress outweighed the risk of misuse by terrorists .

Following the declarations, reports and studies stressing the need to foster open access to publicly funded research results, several important research funding bodies have established or announced policies supporting open access to the research output they have funded (US National Institute of Health, UK Research Councils, Wellcome Trust, CERN, CNRS, DFG, Max Planck Society...). They urge their researchers to publish in open access journals, offering to pay the publication fees, and/or to deposit their articles in an open access repository.

The major public-access concerns are encapsulated by the four fundamental principles on which the UK Research Councils have founded their recent “Statement on Access to Research Outputs”:

- *“Ideas and knowledge derived from publicly-funded research must be made available and accessible for public use, interrogation, and scrutiny, as widely, rapidly and effectively as practicable.*
- *Effective mechanisms are in place to ensure that published research outputs must be subject to rigorous quality assurance, through peer review.*
- *The models and mechanisms for publication and access to research results must be both efficient and cost-effective in the use of public funds.*
- *The outputs from current and future research must be preserved and remain accessible not only for the next few years but for future generations.”*¹⁹

The recent policies of research funding bodies in favour of self-archiving raise some publishers associations’ concerns about the potential negative impact on the viability of their journals.

Although publishers agree with the UK Research Councils’ fundamental principles, they fear that as articles become freely available in open archives and as search, access and retrieval facilities are enhanced by search engines and interoperability, journal subscriptions will be cancelled, thereby undermining the viability of their journals. This might, in turn, decrease the profitability of commercial publishers and reduce the surplus available for learned societies (which they use to cross-subsidize other scholarly activities)²⁰, and thus threaten their existence and activities.

The study commissioned by DG Research starts from this global context of thought process about research policies, in which various stakeholders have been quite active. Its objective is to provide an independent analysis of the conditions regarding access and dissemination of research results, with a view to maximize social returns on R&D investments.

² The most famous disciplinary archive, arXiv, was established in 1991 to provide quick and centralised access to pre-prints in physics and mathematics and improve scholar-to-scholar distribution.

³ As this letter did not invoke the desired response from publishers, PLoS began to publish its own open access journals in 2004, PLoS Biology and PLoS Medicine.

⁴ <http://www.soros.org/openaccess/>

⁵ <http://www.earlham.edu/~peters/fos/bethesda.htm>

⁶ <http://www.zim.mpg.de/openaccess-berlin/index.html>

⁷ SQW Ltd (2003, 2004)

⁸ WSIS (2003).

⁹ “UNESCO between two Phases of the World Summit on the Information Society”, Saint-Petersbourg, 17-19 May 2005. Final document URL: <http://www.wsis-si.org/unesco-russia05-recomm.html> esp. Section 5.

¹⁰ OECD (2004)

¹¹ OECD (2005)

¹² Informally called the “Sabo bill” after the Representative Martin Sabo. Sabo, M. O., <http://www.house.gov/sabo/>

¹³ <http://www.taxpayeraccess.org/>

¹⁴ House of Commons (2004)

¹⁵ Ministry of Education Finland (2005)

¹⁶ Hauge (2005) and Winkel Schwarz (2004-05)

¹⁷ OECD (2004)

¹⁸ National Research Council report (2004)

¹⁹ RCUK (2005)

²⁰ See STM reaction to NIH policy “STM Position on NIH Open Access Proposal” 12 November 2004 (URL: <http://www.stm-assoc.org/statements/accessprop.php>) and ALPSP response to the RCUK position statement (URL: <http://www.alpssp.org/events/2005/RCUKmeeting/RCUKResponse.pdf>).

2. OVERVIEW OF THE MARKET FOR SCIENTIFIC PUBLISHING

The point of view taken in this study is similar to that taken when analyzing markets generally: that of both static and *dynamic efficiency*:

1. Static efficiency requires goods and services to be provided as much as possible at a price that insures cost recovery but avoids ‘excessive profits’. Indeed, if prices do embody higher-than-competitive returns, some consumers will end up avoid buying the goods or services even though their willingness to pay actually exceeds the cost of providing these goods or services. This is why ‘market power’ is monitored by public authorities (in their competition policy), because it is indicative of a supplier’s ability to raise prices above costs, while in a perfectly competitive market the high number of actual and/or potential entrants prevent such behavior.
2. Note that, in a market where a large part of the demand is publicly funded, there is an additional desire to avoid excessively high prices, because these have to be funded from scarce tax revenue which typically have to be raised in a distortive fashion. This is another reason why, in natural monopoly contexts, regulatory authorities try to minimize prices subject to allowing producers to break even.²¹ Of course, the scientific publishing market is not monopolistic (there are several big publishers for example, as well as many smaller ones), and nobody is talking about a natural-monopoly-style regulation as in energy, local telephony, or postal services for example. It is, however, a particularly legitimate concern to avoid excessively high prices of scientific journals, in order to foster cost-effective dissemination of publicly-funded knowledge.

3. While static efficiency concerns tend to favor low prices, it is well-known that dynamic efficiency considerations mitigate this view: ever since Schumpeter (1934), one has recognized the role of ex-post profits as an incentive to bring new goods or services to the market. In particular, it is only natural for publishers who create new journals to earn more than competitive profits on successful new journals as a way to compensate for losses made on journals that end up in failure. “How much more”, is a key empirical question. And a relevant dimension of this question concerns the identity of the publishers that do introduce new journals: Part of the market dynamism, and of incentives to lower prices, could naturally come from new entrants, so the questions of ease of entry, and of potential barriers to entry, are very relevant here too.

2.1. MARKET IMPERFECTIONS

The market under consideration is very far away from the ‘ideal perfectly competitive private market’ that has been celebrated ever since Adam Smith (1776). First, there is the importance of public funding of author and referee time, as well as of journal purchases. Second, researchers, the producers of scientific output, are also consumers of scientific information, and the private and social values of publications may differ when the individual researcher behaves as an author or a reader. Third, authors and readers are not in direct contact: the market is intermediated. Authors are in contact with publishers, which, as stressed above, organize the diffusion and the certification of this information often with the free input of many producers/consumers; and these intermediaries contract not only directly with the final reader but also with other intermediaries, whether aggregators or libraries. All these features tend to strongly weaken the price sensitivity of consumers, be they authors or readers.

They mean that the market for scientific publishing is, in terms of distance from idealized perfectly competitive private markets, in a sense closer to the health market than to the market for, say, haircuts.

An additional complication of the scientific publication market stems from network effects that are commonplace in ‘two-sided markets’ (see for instance, Baye and Morgan, 2001, Caillaud and Jullien, 2003, Ellison and Fudenberg, 2003, Rochet and Tirole, 2003): Authors want to publish in journals that attract other good potential authors and therefore high interest from readers; in turn, readers want to read journals where good authors publish. This can lead to virtuous circles: Through the certification process, journals acquire reputation levels, which may make them ‘unavoidable’ for readers and very attractive for authors. These can be called ‘natural barriers to entry’. With the computation of impact factors and the increased reliance of such indicators in academic incentive schemes, the ‘unavoidability’ of highly-cited journals has been strengthened in recent years. It is important, if one wants to prevent markets from being captive, to avoid ‘strategic barriers to entry’, i.e. those that would be deliberately created by powerful incumbents, from being erected.

We said above that one should aim to a healthy supply of cost-effective, high-quality journals. In fact, while it is especially crucial for the final *reader* to have access to “high quality” journals only, we may also like research agendas to be sustainable by allowing the publication of “lower” quality research. Indeed, publishing, even in journals which are not necessarily the ‘top’ ones in the field, helps the reputation and career of *authors* since these are often linked to “objective” measures of performance like publications. Since the creation of high quality research is generally cumulative, both for the individual researcher and for a field, publication of articles which may be viewed of lower quality thus helps maintain the ‘momentum’ of research agendas. An analysis of the market for scientific publishing must therefore take into account the joint effects between publication

and certification and the different needs and demands that the same individuals may have as producers or as consumers of scientific information.

2.2. MARKET TRENDS UNTIL THE MID-1990’S

Let us first discuss the market evolution before the ‘transition to electronic publishing’, a phase that started around 1995.²²

The scientific publishing market is a market on which very much has been written²³, and it is beyond this study to offer an exhaustive account of the literature. Instead, we concentrate on discussing the key facts that are relevant for a policy evaluation of the functioning of the scientific publication process:

1. Scientific journals represent an essential channel for the diffusion of scientific knowledge, and their importance is, if anything, growing for the research community, in academia and beyond.
2. The ‘business model’ that has dominated the market is the ‘reader-pay model’ (which might be better called, as we shall see below, the ‘library-pay model’), whereby authors do not pay for publishing their paper, possibly beyond an often-symbolic submission fee. Note however that, in this model, authors do not get paid directly by publishers either, and neither do referees, except symbolically again: Authors get rewarded indirectly, in terms of enhanced scientific reputation and its associated career rewards; and referees fulfill their evaluating role out of ‘scientific community norms’. Publishers incur the costs of processing the article (including the associated editorial fees) and of producing and distributing it. This production process involves significant fixed costs (also called ‘first-copy costs’)²⁴. Publishers cover their cost (and possibly more) by selling journals to individual researchers and especially to libraries. Note that in this business model the certification and dissemination functions of scientific knowledge are bundled.

3. There is a large number of active scientific publishers in the market. These can be divided into two big groups: commercial, or for-profit (FP), publishers and not-for-profit (NFP) publishers. NFP publishers include learned societies and university presses. While FP publishers have arrived on the market more recently – i.e. after World War II – on average, their market share has been growing, especially for a couple of big publishers, who have experienced a dramatic increase in their portfolio of journals since the 1970's, because, as stressed by Tenopir and King (2000, page 28): (i) they were the ones starting most of the new journals; (ii) some learned societies turned their journals to them; and (iii) because they grew through several mergers²⁵. Note that, in the aggregate though, the number of journal articles published has been roughly proportional to the growth in the number of scientists (Mabe and Amin, 2001). And since the number of articles published by each journal on average has been on the increase over time, the number of journals has tended to grow less fast in percentage terms than the number of scientists.
4. While the supply of journals seems to have been satisfactory in the decades before the electronic transition, the 1975-1995 period has been dubbed the 'serials crisis', especially because of the dramatic increase in prices of journals sold by FP publishers: an increase of more than 300% *beyond inflation*, and this at a time that preceded the big technological change of the electronic transition. Data collected by Tenopir and King(2000) indicate that this price trend was roughly similar for learned societies – almost a 300% increase beyond inflation too – but from a much lower initial price level. As for university presses, the price increase was more moderate (but still of the order of 200% beyond inflation) and from an even lower initial price level²⁶.
5. Unsurprisingly, as documented for example by Tenopir and King (2000)²⁷, this price evolution, was accompanied by a fall in subscriptions, especially by individual researchers (relying

more on libraries), but also in multiple library subscriptions and in purchases of alternative sources of knowledge like monographs. Indeed, journal prices far outpaced the evolution of library budgets, which in fact did not increase as fast as total academic research budgets. Cummings et al. (1992) show that, in the main research universities of the US²⁸, the libraries' share of general education expenditure rose at the end of the 60s and beginning of the 70s (2.9% to 3.8%) and decreased in the 80s (3.8% to 3.1%). They further report that the share of libraries expenditures on material and binding has been approximately constant (oscillating between 33% and 35%).

The situation until 1995 is therefore one which has a bright side – the centrality of academic journals in scientific exchange and a supply of journals that follows the growth of research – as well as a dark side – dramatic price increases since the 1970's and evidence of reduced circulation as well as crowding out of some alternatives like monographs.

Questions naturally arise concerning the causes of this price evolution as well as concerning the difference in price levels between learned societies and FP publishers in particular. The often-quoted Tenopir and King (2000) study²⁹ stresses a cost-based explanation for both phenomena: (i) FP journals are more expensive because their circulation is lower on average than learned-society journals; and (ii) lower circulation for all journals led to increases in prices. Of course, they recognize the potential reverse causality between prices and circulation. In the next sections of this report, we present evidence that leads us to cast doubt on this cost-based pricing hypothesis. As illustrated in the next section, our evidence is more consistent with the work of McCabe (2002, 2004)³⁰, who rather stresses the positive correlation between journal prices and their 'usefulness' for libraries, i.e. for researchers.³¹ This is in turn consistent with the existence of market power, that is, of the ability of producers to sustain prices in excess of costs.

2.3. TRENDS SINCE THE MID-1990'S

The last ten years have witnessed major technological changes in scientific publishing, with the emergence of electronic publishing. We shall discuss some of these points in detail below. Let us simply stress here a number of important trends:

1. The technology has reduced the marginal cost of dissemination of existing content. On the other hand, there is a one-off cost of making previous content digitally available. And new opportunities to offer new services, like citation data, cross-references, ... Once again, providing these services entails a significant fixed cost but low marginal costs. In this sense, new technologies increase returns to scale, even though it should be stressed that the fixed cost of starting a new journal and making it available on the web has gone down.
2. The importance of scientific journals has further increased. A contributing factor is the fact that access has significantly improved, in that researchers can have instant access, from their computers, to journal content (if it is bought by their libraries!)
3. The business model remains overwhelmingly reader/library-pay, even if experimentation takes place with alternative models, and in particular author-pay, or open access models³² (see House of Commons, 2004, SQW Ltd, 2003, 2004, McCabe and Snyder, 2004, Morris, 2005).
4. In terms of the population of publishers, previous trends continue, the size of the largest publishers increases, mainly through mergers and acquisitions (for a review of the recent mergers and acquisition, see Simba, 2004). Entry of small publishers is also taking place, taking advantage of the lower fixed cost of starting a journal with digital technologies.

5. Prices have kept increasing faster than inflation, even though the rate of increase beyond inflation has slowed down in comparison with the previous 20 years³³. One reason is clearly the limits of library budgets and the pressure that previous price hikes have put on these budgets. This being said, it can be argued that, thanks to the progress in terms of access and in contrast to the previous twenty years, the last ten years have witnessed 'higher value for money' for researchers³⁴.
6. Significant changes have occurred in pricing policies: One has moved very much from *individual pricing* of journals, with one price per journal for individuals and one price per journal for institutions to 'Big Deals', or prices for *bundles of journals that vary from institution to institution*. Moreover, libraries have banded together in 'consortia' in order to share to some extent the benefits of access (and possibly to improve their bargaining positions in front of publishers). All this has made pricing much less transparent than before.

The bright and dark sides discussed earlier concerning the 1975-1995 period are still there: Even if one can clearly argue that new technologies have provided higher value for money for researchers in the last ten years, the legacy of the previous price increases is still very much there, and prices have kept increasing faster than inflation.

Moreover, serious questions arise concerning the role of new pricing practices on the evolution of competition in this market. Indeed, new technologies have opened the door for new business models, lowered some of the costs of production and diffusion, and they can potentially facilitate the entry of new players in the market. In the absence of market imperfections, the market would adjust efficiently to the new technologies: business models that are not profitable should disappear, those that provide superior value to consumers should survive.

However, as stressed before, the publishing market is full of imperfections, be they entry barriers, the fact that the price paid to some of the inputs (say author or referee time) does not reflect their social value, the fact that researchers do not internalize the costs they impose on editors, referees and publishers when submitting their work, or the fact that libraries have only an imperfect knowledge of the value of the journals for the end users. Moreover the feasibility of some business models depends on changes outside the publishing industry: for instance, unbundling certification and diffusion requires that promotion committees start using other instruments than the number of journal articles, open archives require financing from the public sector, an author-pay system requires a change in the way research budgets are allocated, etc.²⁵

When markets are imperfect, some players in the market may be able to influence the dynamic of the industry, and one objective of this report is to highlight the tradeoffs induced by some of these strategies and also to understand how public intervention may facilitate or hinder these dynamics.

²¹ See for example Laffont and Tirole (1993).

²² This is for example the year during which Reed Elsevier, the largest publisher, started investing significant funds on this transition)

²³ See for instance Tenopir and King (2000), the multiple reports commissioned by funding bodies such as the UK House of Commons (2004), the Andrew Mellon Foundation (Cummings et al., 1992), the Wellcome Trust (SQW Ltd, 2003, 2004), by associations of libraries such as the ARL (Kyrillidou and Young, 2004), JISC (Rightscom Ltd, 2005), by associations of publishers or publishers such as ALPSP, Oxford University Press (White and Creaser, 2004).

²⁴ Many authors have emphasized the consequences of high “first copy costs”; see for instance Tenopir and King (2000), Ordover and Willig (1978), Noll and Steimuller (1992).

²⁵ See McCabe (2002), Susman and Carter (2003). Some of these mergers and acquisition have come under the scrutiny of competition authorities; see Competition Commission (2001), DG Comp (2003).

²⁶ This serial crisis has been widely studied. Cummings et al. (1992) report that the average price of periodicals increased at an average annual rate more than one and one-half times that of hardbound books between 1963 and 1990 (11.3% per year Vs 7.2% per year). Moreover, the prices of hardbound books and serials increased more rapidly than the general price level (6.1%). The comparison with hardbound books is interesting because serials and hardbound books share some common production processes.

²⁷ See Ordover and Willig (1978) for a pricing model of individual and library subscription.

²⁸ For other countries see Sowden (2004) [UK + international comparison of 2003 spending], Houghton (2002) [Australia], Obe et al. (2000) [European Union; reports an increase of 33% of expenditures on materials between 1991 and 1998] and Fuegi and Jennings (2004) [updated version].

²⁹ The same kind of explanations have been put forward by Chressanthis and Chressanthis (1994).

³⁰ See also Cummings et al. (1992), Noll and Steimuller (1992), Chressanthis and Chressanthis (1994) for earlier studies.

³¹ This is also implicitly acknowledged by Dirk de Heer (Usage Research Manager of Elsevier): “Keeping a close watch on usage figures of their customers is par for the course for Elsevier sales and account development teams. It’s in the mutual benefit of customers as well as Elsevier for usage to be sufficiently high compared to consumers spend.” See Dirk de Heer (2005), How Elsevier Uses Usage Information: The Inside Scoop, Library Connect, Elsevier. (http://www.elsevier.com/framework_librarians/LibraryConnect/lcpamphlet7.pdf)

³² For a good description of new business models as well as current experiments, see Rightscom (2005).

³³ Kyrillidou and Young (2004), page 11, report an increase of the serial unit cost 140% between 1986 and 1995, well above 32% between 1995 and 2003 or, in real terms: 73% and 9% respectively. Similarly, SQW Ltd (2003) reports a price increase of 80% between 1990 and 1995 above 55% between 1995 and 2000, or in real terms: 52% and 36% respectively. See also CSFB (2004).

³⁴ House of Commons (2004) pp 34-35, Kyrillidou and Young (2004), Nicholas et al. (2005), Ball and Pye (2005).

³⁵ For instance, at ICOLC 2004 (<http://www.collectionscanada.ca/consortium/s38-1012-f.html>), Springer indicates that it offers authors an “Open Choice”: the author can make his or her published article openly available. The cost of this option for the author is \$3,000. Oxford University Press indicates a price of \$1,500 per paper for full open access.

3. ENTRY, PRICES, COSTS AND QUALITY

While particular in many respects, the performance of the market for scientific publishing can be evaluated with respect to the same variables used in other industries: price, quality and entry.

From a *static perspective*, that is when one takes a short-run perspective, both the price and the “quality” are important variables.³⁶ Everything else being equal, lower prices are socially beneficial since more consumers can access the product and since the price becomes closer to the social cost of production. Everything else being equal, higher “quality” is also a good thing from a social point of view.

Quality can be defined in one of two ways. A first concept of quality is the usual concept of “being better”: end users prefer to read important and well crafted papers, researchers prefer to be published in journals carrying a good reputation. Industrial organization refers to this concept as a concept of “vertical differentiation”. A second concept of quality is related to the view that consumers like diversity and value market offerings with many varieties (e.g., many journals): end users because knowledge is cumulative and different papers can be complementary, researchers because more journals, even if differentiated on the basis of reputation, increase the probability of publication and therefore of meeting basic criteria for promotion. This view is central in the literature on monopolistic competition in industrial organization and more varieties bring social value because consumers are “horizontally differentiated”.

Higher prices are no longer inconsistent with a larger social surplus if it is accompanied by a larger quality, whatever the quality concept used.

However, the question is whether prices could decrease, or whether a price differential between two products reflects differences in costs. In the vertical definition, this requires that higher quality increases the cost of production. While it seems rather natural that higher quality would come at a higher cost, this assumption is not necessarily pertinent for scientific publishing. In the horizontal definition, a publisher distributing many journals in a given field – and therefore increasing the “horizontal” quality – will have larger costs only if there are diseconomies of scale, something that most publishers will view as quite unrealistic.

Hence from a static perspective, potential inefficiencies come from price differentials between journals or offerings of different quality that do not reflect differences in costs.

As far as costs are concerned, it has often been argued (see for example Tenopir and King (2000)) that one of the causes for the large increases in prices (besides inflation, inflationary cost of labor, fluctuation in exchange rates, ...) are due to the decrease in circulation. For profit publishers also argue that their journals are more costly because, on average, they produce more journals serving “niche” audiences and fewer journals dedicated to a larger community of scientists.

In order to understand whether differences in prices reflect differences in costs, we need to have an idea of the differences in costs per subscription of an article. There is a wide literature on publishing costs (Tenopir and King, 2000, Dryburgh Associates Ltd for ALPSP, 2002, SQW Ltd for Wellcome Trust, 2004, Marks, 1995, Morris, 2005). These publications are mainly concerned with the different costs incurred to produce an article. The authors emphasize the importance of the cost of the “first copy”, and show that, for a constant level of quality, the cost per subscription of an article decreases steeply when the number of subscriptions increases. To give orders of magnitude, following calculations made at the end of the 1990’s by Tenopir and King (2000), recovery costs decrease from around \$775 per subscription for 500 subscribers to around \$107

per subscription for 5000 subscribers (86% decrease; see Tenopir and King (2000), figure 11 page 266).

However, the question we will deal with here is slightly different, since we will let quality vary. Indeed, the data we have concern total citations rather than circulation or costs, which are very hard to obtain. But looking at the interaction between prices and citations can in our view be instructive.

Higher citations go hand-in-hand with higher circulation, which should reduce costs, but there is also a counterbalancing effect on costs, for example because higher-quality journals attract more submissions and have a higher rejection rate. Which effect should be expected to dominate?

It is in fact possible to have an idea of the proportion of the total costs of an article that will increase with the level of quality for a given domain. Using the methodology of Tenopir and King (2000) it is clear that some costs (e.g. refereeing) may increase while other costs should not vary much (e.g. print). Specifically, Tenopir and King (2000) distinguish Article processing costs, Non-article processing costs, Reproduction costs, Distribution costs and Publishing support costs. There is no reason to believe that the Publishing support (general administration, taxes), Reproduction or Distribution costs vary much with quality (as measured by citations for instance) of a journal in a given domain. The same holds for Non-article processing costs (table of contents, letters, editorials), which should not be correlated with quality.

Some of the two remaining costs may be related to quality. Article processing costs include five sub costs, two of which may change substantially with quality of the review:

- Fixed costs per issue, which should not change much;
- Cost per page of receiving, processing, and reviewing a manuscript: these can increase substantially with the level of quality and in particular with the rejection rate. Rejection rates vary much more across domains than within a given domain but increase with quality;

- Costs of editing and proof-reading articles may also increase with the quality of the journal;
- Costs of composition and processing graphics should be quite stable;

The costs, which vary with quality, thus represent, following the estimates of Tenopir and King (2000), a bit more than 60% of the Article costs, which are themselves one third of the costs of an article. This means that some twenty percent of the costs of producing an article vary with quality. This figure may seem low compared with the steep decrease of cost per subscription with the number of subscriptions. Moreover, apart from costs, higher quality journals have higher incomes. Besides the higher income from subscription, they may charge for advertisement and sell more back issues and reprints. These additional incomes may be nontrivial.

Average costs should therefore be expected to fall when citation counts rise³⁷.

In the remainder of this section, we investigate empirically the relation between journal prices, their for-profit or not-for-profit status and their quality measured by their citation counts. We do this on a large sample of journals covering 22 scientific fields.

3.1. DESCRIPTION OF THE DATABASE

Many authors have provided evidence indicating that journals held by FP publishers are significantly more expensive than journals published by NFP organizations. However, these studies are either concentrated on a precise domain³⁸ or aggregate many domains together³⁹. As will be clear later on, there are wide differences across domains. This makes an analysis on a precise domain not very relevant for another domain and may introduce a bias in studies that aggregate too many domains together (an unbalanced database may introduce biases). This has led us to study journal prices in 22 domains of natural and social sciences.

We have selected the domains in the following way:

- “domains” are defined by the *Journal Citation Reports* (Journal Citation Reports, 2003—JCR in what follows)
- all domains with more than 135 journals were selected in science (11 domains) and all domains with more than 80 journals in the social sciences (6 domains). This selection criterion unfortunately excluded chemistry and physics, which are subdivided into quite small domains, with the consequence that no domain contains 135 journals. We therefore included five more domains, which did not match the 135-journals criterion in science, but were the five domains with the most journals in chemistry and physics.

The 22 domains studied are:

- for the sciences: 1. Biochemistry and Molecular Biology; 2. Cell biology; 3. Chemistry, multidisciplinary; 4. Chemistry physical; 5. Clinical neurology; 8. Engineering, chemical; 9. Engineering, electrical and electronic; 11. Materials science, multidisciplinary; 12. Mathematics applied; 13. Mathematics; 14. Neuroscience; 15. Pharmacy; 16. Physics, applied; 17. Physics, multidisciplinary; 18. Plant science; and 22. Surgery.
- for the social sciences: 6. Economics; 7. Education and educational research; 10. Law; 19. Psychology, clinical; 20. Psychology, multidisciplinary; and 21. Sociology.

The number of journals in the different domains is reported in column 1 of Table 1. These domains are not mutually exclusive: The same journal can be classified in “Mathematics, applied” and in “Mathematics” for instance. However, less than ten percent of the journals in the domains that we selected appear in more than one domain.

We then classified the journals into two categories: (a) *NFP journals* published by not-for-profit publishers (scientific societies, University presses, etc.) including those published by for-profit firms on account of scientific societies and (b) *FP journals* which belong to for-profit societies.

To do this we looked for the journals on the World Wide Web. Many journals can be classified easily in one category, but for some journals it is less so. We considered that a journal published by a for profit publisher on behalf of a not-for-profit organization is not-for-profit. In a small number of cases we had to make choices, some of which were subject to some discretion.⁴⁰

SWETS provides a database that includes in particular the following information for approximately 130.000 journals: prices (domestic and foreign between 2001 and 2004), publisher, ISSN number, date of first publication. The JCR (2003) provides the number of citations. The two databases were merged. Some journals appearing in the JCR are not sold by SWETS. However, there was no systematic bias (FP and NFP publishers as well as large and small publishers are both represented in the database).

The SWETS database provides eight entries for prices: Domestic and foreign, each for four years. These prices are expressed in various currencies (usually the local one, that is the currency of the country in which the publisher is located). We converted these prices into Euros using exchange rates from IMF (World Economic Outlook, 2005). The evolution of prices cannot be used directly for at least two major reasons: (a) the Euro was introduced in 2001 and several prices are still expressed in national currencies in 2001 (and even in 2002). Note that some publishers significantly increased their prices at that moment; and (b) the Euro has appreciated by some 40 percent against the dollar between July 2001 ($\$1 = 0.87$ Euro) and July 2004 ($\$1 = 1.21$ Euro), that is during the period under study, an effect that cannot be corrected for.⁴¹ This may be at the origin of a systematic bias that makes NFP journals, mainly located in the United States, look cheaper than FP journals, more frequently located in Europe. However, this bias should not be large, since we take an average of prices over the whole period, which corrects for the two extreme values.

Moreover, some prices were missing. In order to circumvent the various problems, we created a price indicator which is the average of available prices, but there are both NFP and FP journals which had to be dropped from the analysis, because we could not retrieve their prices (or could not construct a convincing indicator).

The original database did not give much of the information on creation dates of journals (less than 50%). The missing information was collected from the web (libraries and groups of libraries give this information).⁴²

The total number of journals for which we could retrieve the full set of information that will be used is 2,707, and is very likely to be representative for top journals, that is, those which are part of ISI, and which probably receive the largest number of citations. These are also the journals, which are read the most by scientists, even if they also publish in other journals.

TABLE 1 GENERAL STATISTICS: PRICES, CITES AND AGE

	No. of journals	NFP journals (%)	Prices					Cites/journal					Age							
			FP		NFP		R	S	FP		NFP		R	S	FP		NFP		R	S
1. Biochemistry and molecular biology	261 (213)	18	1343	768	1.75	**	5863	20330	0.29	23	29	0.79	*							
2. Cell biology	156 (134)	27	1270	556	2.29	**	6172	7746	0.80	22	26	0.83								
3. Chemical, multidisciplinary	123 (101)	55	975	566	1.72	*	1945	8952	0.22	29	43	0.68	**							
4. Chemistry, physical	101 (87)	19	2389	1334	1.79	*	3737	7187	0.52	29	25	1.18								
5. Clinical neurology	134 (126)	19	646	375	1.72	*	3036	4370	0.69	26	33	0.80								
6. Economics	169 (151)	40	613	188	3.26	**	590	1231	0.48	30	47	0.63	**							
7. Education and educational research	92 (86)	45	523	158	3.30	**	275	366	0.75	30	46	0.66	**							
8. Engineering, chemical	119 (100)	27	1359	458	2.97	**	1540	1802	0.85	30	36	0.85								
9. Engineering, electrical and electronic	205 (172)	54	1049	522	2.01	**	595	2829	0.21	25	28	0.90								
10. Law	102 (101)	78	385	68	5.63	**	423	623	0.68	33	51	0.64	**							
11. Materials sciences, multidisciplinary	177 (147)	22	1569	603	2.60	**	2211	2697	0.82	22	22	1.03								
12. Mathematics, applied	153 (147)	20	1344	468	2.87	**	879	1472	0.60	25	30	0.84								
13. Mathematics	174 (157)	36	1104	340	3.25	**	863	980	0.88	36	51	0.70	**							
14. Neuroscience	198 (182)	18	1322	582	2.27	*	4278	6958	0.61	24	31	0.79								
15. Pharmacy	185 (168)	19	1250	409	3.06	**	2671	4876	0.55	24	39	0.62	**							
16. Physics, applied	76 (69)	40	1791	1150	1.56		1979	11073	0.18	24	31	0.78								
17. Physics, multidisciplinary	68 (57)	50	1735	1060	1.64		4327	9931	0.44	45	39	1.16								
18. Plant science	136 (123)	49	1038	398	2.61	**	2111	3266	0.65	33	54	0.62	**							
19. Psychology, clinical	83 (79)	33	458	207	2.21	**	802	2349	0.34	25	34	0.73	*							
20. Psychology, multidisciplinary	101 (92)	41	444	179	2.48	**	597	1766	0.34	27	46	0.58	**							
21. Sociology	93 (87)	47	367	145	2.52	**	278	734	0.38	30	47	0.65	**							
22. Surgery	141 (128)	17	516	270	1.91	**	3658	3846	0.95	32	48	0.67	**							

Notes. FP = for profit; NFP = not for profit; R = ratio FP/NFP; S = significance level for testing that the means are equal; ** (or *) = equality of means rejected at a probability level that is smaller than 1 % (or 5%); in other cases, equality of means cannot be rejected (level larger than 5 %). No. of journals: the first number refers to journals in ISI, the second to those for which we could retrieve all the information (prices, cites and age).

3.2. DESCRIPTIVE STATISTICS

Table 1 reports the main characteristics of the journals in each domain: percentage of NFP journals (column 3) and, for each type of publisher (FP and NFP) and each domain, average prices, average number of citations per journal, and average age (2005 minus date of first issue). The columns entitled “R” show the ratios FP to NFP for prices, citations and age. As can be checked, *for every single domain*, FP charge higher prices than NFP journals, and have fewer citations. They are, on average, older in only two domains (Chemistry, physical; Physics, multidisciplinary). This provides a clear confirmation of the more scattered (and scarce) information discussed in previous studies. Differences between domains are large, and sometimes larger than differences between NFP and FP journals within the same domain.

The citation pattern and the date of creation of journals indicates that FP publishers have been faster to expand output (they have created many titles with few cites, especially recently). The expansion of FP publisher outlets has provided the supply to meet the development of research output. Indeed, several studies (e.g. Tenopir and King (2000), Mabe and Amin, 2001) indicate that the creation of journals has broadly followed the growth in the number of researchers.

Note also that there is heterogeneity within categories. This is visible when we compare the different domains. Online journal creation has been embraced by some scientific societies but not so much by others. For instance, the American Economic Association has only three journals, both available in print and digital formats, the American Psychological Society has 49 journals; the engineers' society IEEE has 126 journals, and sells bundles, both in print format or online, the Society for Industrial and Applied Mathematics (SIAM) has 13 online journals, the American Chemical Society has 33 journals in the ISI-JCR and houses abstracting and indexing collections (the Chemical Abstracts Service) while the American Medical Association seems to have only one, the celebrated *JAMA*, but also has *JAMA & Archives* for a very large number of sub fields, though it is not clear whether these are all separate journals.

The Econometric Society which launched *Econometrica* in 1933 never launched another journal. It just decided to start two new electronic journals.

The picture that emerges is thus the following. At one extreme, as stressed by Tenopir and King (2000), some learned societies manage a small number of highly successful long-standing journals, the best in the profession, which they sell for low prices (but without making big losses⁴³). They serve the ‘high end of the market’, and seem happy to keep doing so. At the other extreme, for-profit publishers are introducing new journals, accompanying the growth in research output. On average, these for-profit journals are of lower quality, as measured by cites, but they are sold at much higher prices.

Economics fits this pattern. While we have stressed that FP journals have a lower average quality than NFP journals, there is however quite a lot of heterogeneity in these two categories. In particular, while the best ‘general’ journals are non-profit (e.g. the top-five economics journals: *American Economic Review*, *Econometrica*, *Journal of Political Economy*, *Quarterly Journal of Economics* and *Review of Economic Studies*), a majority of ‘top field journals’ are for-profit (e.g. *Journal of Development Economics*, *Journal of Monetary Economics*, *Journal of Public Economics*, *Journal of International Economics*, *Journal of Econometrics*, *Journal of Economic Theory*, ...). These are journals that have been created in the 1960’s and 1970’s, when the expansion of research output led to the increasing specialization of the research community, which in turn led researchers to prefer a publication in a top field journal than in a second-tier general journal. The pace of creation of new field journals has clearly been much faster than the emergence of specialized learned societies, even though this creation has been happening too.

Of course, these broad tendencies should be refined. It would, for example, be interesting to measure the typical price difference between an FP and an NFP journal, for a given field and a given journal quality (measured by the total number of cites). This is what we turn to now.

3.3. PRICES AND QUALITY: AN ECONOMETRIC ANALYSIS

3.3.1. A FIRST, SIMPLE ANALYSIS OF PRICES, CITATIONS AND AGES

Consider first simple comparisons of average prices. The columns entitled “S” in Tables 1 and 2 give the results of statistical tests to check whether the averages obtained for FP and NFP journals are statistically different (the usual t-test). These show that the null hypothesis that averages are equal is rejected in 20 out of the 22 cases for prices (at a smaller than 1% probability level in 16 cases, see **, at a smaller than 5% level in the four remaining cases, see *). Thus FP journals are almost always more expensive than NFP journals. The same test shows that this is also so for citations in 9 cases (at a smaller than 1% probability level in 5 cases). FP journals are less cited than NFP journals. Finally, NFP journals are significantly older than FP journals in a little more than half of the cases. Equality could not be rejected in the remaining 10 cases.

Since some journals carry more articles than others, we also compared prices per article. This is detailed in Table 2, where we repeat average prices, and compare them with prices per article. Again, FP journals are systematically more expensive than NFP journals, and significantly so in 18 cases out of 22.

TABLE 2 GENERAL STATISTICS: PRICES AND PRICES PER ARTICLE COMPARED

	No. of journals	NFP journals (%)	Prices			Prices/article				
			FP	NFP	R	S	FP	NFP	R	S
1. Biochemistry and molecular biology	213 (209)	18	1343	768	1.75	**	14.78	5.40	2.74	**
2. Cell biology	134 (133)	27	1270	556	2.29	**	16.54	6.99	2.37	**
3. Chemical, multidisciplinary	101 (99)	55	975	566	1.72	*	10.76	5.74	1.87	**
4. Chemistry, physical	87 (85)	19	2389	1334	1.79	*	22.00	14.25	1.54	
5. Clinical neurology	126 (121)	19	646	375	1.72	*	8.39	5.79	1.45	
6. Economics	151 (150)	40	613	188	3.26	**	15.14	6.91	2.19	**
7. Education and educational research	86 (82)	45	523	158	3.30	**	16.74	6.10	2.75	**
8. Engineering, chemical	100 (98)	27	1359	458	2.97	**	14.68	4.83	3.04	**
9. Engineering, electrical and electronic	172 (170)	54	1049	522	2.01	**	16.91	5.19	3.26	**
10. Law	101 (99)	78	385	68	5.63	**	14.24	3.25	4.38	**
11. Materials sciences, multidisciplinary	147 (144)	22	1509	603	2.60	**	17.52	5.45	3.21	**
12. Mathematics, applied	147 (144)	20	1344	468	2.87	**	24.21	10.81	2.24	
13. Mathematics	157 (153)	36	1104	340	3.25	**	15.31	8.25	1.86	**
14. Neuroscience	182 (178)	18	1322	582	2.27	*	14.69	6.47	2.27	**
15. Pharmacy	168 (168)	19	1250	409	3.06	**	15.92	4.01	3.97	**
16. Physics, applied	69 (69)	40	1791	1150	1.56		14.10	4.32	3.26	**
17. Physics, multidisciplinary	57 (56)	50	1735	1060	1.64		17.31	12.97	1.33	
18. Plant science	123 (120)	49	1038	398	2.61	**	11.19	4.74	2.36	**
19. Psychology, clinical	79 (77)	33	458	207	2.21	**	12.97	5.04	2.57	**
20. Psychology, multidisciplinary	92 (91)	41	444	179	2.48	**	14.08	5.74	2.45	**
21. Sociology	87 (78)	47	367	145	2.52	**	18.20	6.61	2.75	**
22. Surgery	128 (127)	17	516	270	1.91	**	5.39	2.93	1.84	*

Notes. FP = for profit; NFP = not for profit; R = ratio FP/NFP; S = significance level for testing that the means are equal; ** (or *) = equality of means rejected at a probability level that is smaller than 1% (or 5%); in other cases, equality of means cannot be rejected (level larger than 5%). No. of journals: the first number refers to journals for which we could retrieve all the information (prices, cites and age), the second (between brackets) to those for which we could compute the price per article.

These statistics also give an indication concerning the degree of concentration. In Table 3, we have computed the shares of the main not-for-profit (scientific societies, University presses) and for-profit firms in the 22 domains that were described earlier. In this case, journals published by a for-profit firm on account of a not-for-profit firm is included under the name of the for-profit firm. In industry concentration studies, shares are usually expressed in terms of turnover (here, prices multiplied by circulation of the journal) or profits for each firm. None of these numbers could be computed in our case, since we do not have access to circulation figures, and profits per domain are of course unavailable, and can probably not even be computed by the firms themselves.

We had therefore to base our calculations on the data that were available, that is number of journals or citations. We thought that the number of journals owned by a firm is not a good indicator, since some journals are well known and read, others are much less so. The alternative is the number of citations, which seems to be a good proxy for importance and circulation of the journal.

Table 3 gives an overview of the "market shares" in terms of citations in the 22 domains. We singled out the most important publishers, that is those which are involved in almost all the domains: Blackwell (14 domains); Elsevier (22 domains); John Wiley (21 domains); Kluwer (21 domains); Lippincott (4 domains only, but very heavily in 2 domains); Springer (16 domains); Taylor & Francis (21 domains). These are all for-profit firms some of which may manage NFP journals. We also give the share of the American Scientific Society publishers (Am. Sty in the table) in the domains where there exists one or several such societies (in which case a *is added), plus a few extra important not-for-profit and for-profit firms (under Other in the table).

As can be seen, Elsevier is, the most important for-profit publisher in terms of citations, but its market share is close to or larger than 50 percent in three domains only: 56 percent in Chemistry, physical, 57 percent in Engineering, chemical, 49 percent in Materials sciences, multidisciplinary. Its market share is also quite important in several other domains, such as Economics (29 percent), Mathematics, applied (37 percent), Neuroscience (35 percent); it is lower than 29 percent in all other domains.

Blackwell is reasonably important in Economics, but its market share is still smaller than 30 percent, and Lippincott reaches 32 percent in Clinical neurology.

³⁶ From a dynamic perspective, it is important to assess whether incumbents can leverage their position to increase entry barriers in the industry, leading to fewer entrants or to the non-adoption of potentially beneficial business models. This issue will be discussed in a later section of this report.

³⁷ See also Cummings et al. (1992) or Noll and Steimuller (1992).

³⁸ Chressanthis and Chressanthis (2004) studied economics, Bergstrom (2001) economics (extended by Bergstrom and Bergstrom (2004) to 5 other fields), Binman et al. (1997) mathematics, Cornell (1998) agriculture and biology.

³⁹ McCabe (2002), Tenopir and King (2000), Kyrillidou and Young (2004), White and Creaser (2004), SQW Ltd (2003).

⁴⁰ Note that, to the extent that some journals published by a FP publisher on behalf of a NFP organization is run 'partly' as a FP journal, this leads to underestimate the extent to which FP journals are more expensive than NFP journals, thus amplifying our result below.

⁴¹ Note however, that for British journals, exchange rates fluctuated much less.

⁴² We are grateful to Elisabetta Lazzaro for her help in this data-gathering exercise.

⁴³ Usually publishing is used to cross-subsidize other activities, not the reverse. Baldwin (2004) shows that 2/3 of the scientific societies surveyed use the surplus of publishing activities to subsidize other activities ranging from annual conference to research grants and public education. See also Morris (2005).

TABLE 3 CITATION MARKET SHARES FOR MAIN NFP AND FP (%)

	Blackwell	Elsevier	J. Wiley	Kluwer	Lipp.	Springer	T & F	A. m. Sty	Other
1. Biochemistry and molecular biology	4	23	2	2	-	2	1	26*	10 (Cell Press)
2. Cell biology	2	14	3	2	-	3	1	13*	18 (Cell Press)
3. Chemical, multidisciplinary	ng	3	18	3	-	ng	1	49	10 (Roy. Sty. Ch.)
4. Chemistry, physical	-	56	4	5	-	2	1	23*	5 (Roy. Sty. Ch.)
5. Clinical neurology	7	14	8	1	32	6	1	9*	
6. Economics	21	29	2	3	-	1	2	13*	9 (U. C Chicago)
7. Education and educational research	8	7	7	4	-	-	12	16*	
8. Engineering, chemical	-	57	14	2	-	-	5	12	
9. Engineering, electrical and electronic	-	11	1	1	-	1	1	-	72 (IEEE)
10. Law	2	1	1	3	-	-	-	-	10 (Harvard Law R. Ass.)
11. Materials sciences, multidisciplinary	-	49	6	6	-	3	3	8	
12. Mathematics, applied	-	37	11	6	-	8	2	6	16 (SIAM)
13. Mathematics	-	23	5	5	-	17	2	9	
14. Neuroscience	7	35	11	1	4	5	1	5*	10 (Soc. Neurosc.)
15. Pharmacy	4	29	4	2	4	6	2	18*	
16. Physics, applied	-	14	ng	2	-	3	1	54	11 (IEEE)
17. Physics, multidisciplinary	-	24	ng	1	-	ng	ng	58	
18. Plant science	15	10	ng	10	-	9	2	24*	
19. Psychology, clinical	1	11	7	9	-	-	6	30	11 (Phys. Postg. Press)
20. Psychology, multidisciplinary	5	5	3	4	5	-	4	48	
21. Sociology	14	4	-	4	-	-	7	15	12 (Sage)
22. Surgery	1	7	6	-	27	9	1	9*	16 (Saunders)
									10 (Mosby)

Notes. Lipp = Lippincott; T & F = Taylor and Francis; Am. Sty = American Scientific Society; * = more than one society; Other = Other important NFP or FP; ng = negligible.

Note also the high market shares of learned societies in some specific fields: 72% in Engineering, electrical and electronic (for IEEE), 49% in ‘chemical multi’, 54% in ‘physics applied’, 58% in ‘physics multidisciplinary’, and 48% in ‘psycho multidisciplinary’.

Note that we are well below the 50 percent share in profits or turnover for each firm globally, but this is also the case in terms of citations in most domains that we have looked at in some detail, with the exception of Elsevier in three domains: Chemistry, physical, Engineering, chemical, and Materials sciences, multidisciplinary.

3.3.2. ON CHARACTERISTICS AND PRICES

The question that we try to tackle now is whether some characteristics seem relevant to publishers, both NFP and FP, when pricing their journals. To examine this issue, we estimate the following model for each of the 22 domains:⁴⁴

$$p_{i,j} = a_0 + b_0 * FP + c_0 * Ncites + c_1 * Ncites * FP + d_0 * Nage + d_1 * Nage * FP + c_1 * E + a_j.$$

In this equation, subscript *i* represents a journal, and *j* a field (*j* = 1, 2, ..., 22). The equation shows that each field is represented by a dummy variable which picks a regression coefficient *a_j*.

- *p* is the log of the price (price) or the log of the price per article (*price/art*);
- *FP* is a dummy variable which takes the value 1 for FP journals;
- *Ncites* is the log of the field-normalized number of citations. Since we pool over fields, and since some fields have many more citations than others, we had to normalize somehow. This is done by dividing the number of cites to journal *i* which belongs to field *j* by the average number of cites for all journals belonging to the same field *j*. *Ncites* should thus be interpreted in relative terms. An increase in this number means that a journal is relatively more cited. When we explain the price per article, we naturally use cites per article. *Ncites* is then the number of

cites per article to journal *i* which belongs to field *j* divided by the average number of cites per article for all journals belonging to the same field *j*;

- *Nage* is the log of the field-normalized age of the journal (2005 minus year of first publication). Since there may be systematic differences between domains, the variable *Nage*, where the age of journal *i* which belongs to field *j*, divided by the average age of all journals belonging to the same field *j*;
- *E* is a dummy variable with value equal to 1 for English speaking journals.
- We allow for interaction terms, to see whether the relation between citations and prices, or age and prices, differ between FP and NFP publishers. So we are not only interested in the average price difference between FP and NFP journals, but also in their relative age and quality sensitivities.

A single equation is estimated using all the observations, whether the journal is published by a scientific society or not.⁴⁵ Note that the regression coefficients for *Ncites* and *Ncites*FP* as well as for *Nage* and *Nage*FP* are common to all fields.

In terms of the signs of the coefficients, the ones involving citations are particularly interesting. Indeed, there are two effects here. When a journal is cited, this is usually taken as an indication of “quality” and of high value to the reader. This can turn a profit-maximizing publisher to raise its price since the willingness to pay of readers increase. McCabe (2002, 2004) takes this approach when analyzing the behavior of FP publishers facing libraries which he argues try and buy the ‘best-value-for-money journals’. Note however that, as discussed earlier in this section, we should expect that more citations come with lower average costs. A publisher wishing simply to balance its budget may then be expected to set lower prices for a more heavily cited journal.

As for the influence of the age of the journal on its price, there are also two effects that may go in opposite directions. When a journal is launched, there are sunk costs that have to be paid.

Tenopir and King (2000) estimate that accumulated profits of a new journal are negative during the first six years. The journal may thus charge a higher price when it is new. On the other hand, the publisher may use an introductory price strategy: low price first to attract new readers and then increase prices once the journal is mature and with a higher profile.

In Table 4, we show the results of two regressions. In the first equation, the dependent variable is the price per journal, in the second, it is the price per article per journal.

Since the results are very similar, our comments concentrate on the first equation. We have run the same regressions for individual domains and the results are globally similar.

The results lead to the following comments. The overall results are good in a statistical sense: All the estimated parameters⁴⁶ are very significantly different from zero at the less than one percent probability level, and the fit, measured by the R-square is very satisfactory.

TABLE 4 REGRESSION RESULTS

	Prices			Prices/article		
	Coeff.	Std. err.		Coeff.	Std. err.	
For Profit dummy (FP)	1.10	0.04	**	1.03	0.04	**
No. of Citations (normalized per domain)	0.24	0.02	**	0.14	0.02	**
No. of citations*FP	0.14	0.02	**	0.10	0.03	**
Age (normalized per domain)	-0.18	0.03	**	-0.29	0.04	**
Age*FP	0.14	0.04	**	0.14	0.05	**
English language	0.44	0.06	**	0.27	0.06	**
DOMAINS						
1. Biochemistry and molecular biology	0.95	0.08		-0.31	0.09	
2. Cell biology	0.83	0.08		-0.16	0.10	
3. Chemical, multidisciplinary	0.82	0.09		-0.60	0.11	
4. Chemistry, physical	1.29	0.10		0.11	0.11	
5. Clinical neurology	0.21	0.09		-0.80	0.10	
6. Economics	0.00	-		0.00	-	
7. Education and educational research	-0.38	0.10		-0.12	0.11	
8. Engineering, chemical	0.66	0.09		-0.35	0.11	
9. Engineering, electrical and electronic	0.73	0.08		-0.31	0.09	
10. Law	-1.14	0.09		-0.73	0.11	
11. Materials sciences, multidisciplinary	0.83	0.08		-0.32	0.10	
12. Mathematics, applied	0.66	0.08		0.14	0.10	
13. Mathematics	0.34	0.08		-0.04	0.09	
14. Neuroscience	0.64	0.08		-0.28	0.09	
15. Pharmacy	0.60	0.08		-0.37	0.09	
16. Physics, applied	1.33	0.10		-0.53	0.12	
17. Physics, multidisciplinary	1.28	0.11		-0.18	0.13	
18. Plant science	0.39	0.09		-0.40	0.10	
19. Psychology, clinical	-0.16	0.10		-0.31	0.11	
20. Psychology, multidisciplinary	-0.21	0.09		-0.16	0.11	
21. Sociology	-0.41	0.10		-0.05	0.11	
22. Surgery	-0.01	0.09		-1.34	0.10	
R-square	0.59			0.35		
No. of obs.	2707			2651		

Notes. Coeff = regression coefficient; Std. err. = Standard error; ** means that the coefficient is different from zero at a probability level that is lower than 1%. Probability levels are not mentioned for the various domains, since the standard error merely measures the difference w.r.t. the domain "Economics".

First, there are large differences in prices across domains. Since the control field is Economics, its price is normalized to 100. The "cheapest" field is Law, with a relative price of $100\exp(-1.14) = 32$, the "most expensive" is "Physics, applied" which charges a relative price of $100\exp(1.33) = 378$. Average prices vary thus between 32 and 378, a factor of 1 to 12.

Second, FP journals charge on average $100\exp(1.10) = 300$, that is three times more, as shown by the coefficient β_0 picked by the FP dummy, all other things, i.e. domain, age and total citations being equal. This confirms earlier findings.⁴⁷

Third, journal prices also vary with their number of citations and their age. The price elasticity of a one percent increase in (field-normalized) citations is positive: It is equal to 0.24 for NFP journals, and 0.38 (that is $0.24+0.14$) for FP journals. The price elasticity of a one percent increase in age is negative (-0.17) for NFP journals and close to zero ($-0.04 = -0.17+0.13$) for FP journals. The fact that the coefficient for FP journals is smaller than for NFP journals is probably due to FP journals being born later than NFP journals. As a result, FP journals put less weight on age than NFP journals; but they take advantage to a larger degree of the number of citations that they receive.

Finally, English-speaking journals are 55 percent more expensive than journals in other languages.

We have run several other regressions to check the robustness of our results. One is shown in columns 3 and 4 of Table 4. As was mentioned earlier, results are qualitatively identical. FP journals charge more, and more per cite than NFP journals, though as one could expect, the R-square is smaller. Using Impact factors instead of Cites also yields similar results but lower R-squares.

3.3.3. CONCLUSIONS

Our empirical investigation has documented three facts:

1. FP publishers have given new impetus to scientific publishing (they launched many new field journals) while NFP publishers have focused on high quality journals.
2. There exist large price differences across fields, even taking the number of citations into account. These are as significant in a sense as the differences between for-profit and not-for-profit publishers.
3. As a general rule, FP journals charge more (three times as much on average) than NFP journals, for a given domain, age and citation count.
4. Prices are positively correlated with quality: Both NFP and FP journal prices are positively correlated with the number of citations they receive, and the coefficient is higher for FP journals.

We take the first finding as indicative of the fact that substitution possibilities across journals are limited, allowing for a significant amount of discretion in the setting of journal prices. Indeed, the technology of producing journals is widely available across fields of study, and the financial rewards offered to scientists as authors, referees and editors is sufficiently limited that differences in their labor market opportunities should not translate into very significant journal production costs across disciplines. Our finding is therefore indicative of large differences in price-cost ratios across disciplines.

Our second and third findings are consistent with the first one. We confirm earlier findings concerning the large price difference between FP and NFP journals. Moreover, for each journal category, we show that prices increase with citation counts while we have argued that costs should tend to fall when citation counts rise. This is consistent with 'value-based pricing' (à la McCabe, 2002, 2004) rather than with cost-based pricing, and is again indicative of publishers' ability to exercise discretion in price setting.

Note that these findings do not invalidate the fact that, as stressed for example by Tenopir and King (2000), on *average*, for-profit journals may have higher costs than some learned societies because they have lower average circulations. But our evidence (conclusion number 4) indicates that *individual journal prices do not seem to follow costs*. Pricing is instead consistent with the idea that successful journals are priced higher so as to get a return on previous investment, and in particular so as to recoup the initial investment cost of starting new journals. Unsurprisingly, this feature especially characterizes FP journals. From the point of view of market analysis, this ability of publishers to pursue this pricing strategy is indicative of the power of journals that have managed to establish a reputation of quality, an issue we will return later in this report.

Let us close this section of the report by discussing one aspect of prices and costs that we have not stressed much yet here: the ‘electronic revolution’. Our analysis has concentrated on individual journal prices, that is, the prices of print journals. We have of course entered now the era of at-least-partly-electronic journals and of bundled pricing. Our feeling is however that looking at individual prices for print journals is still quite relevant at this point, because they represent influential base prices that libraries face. Indeed, bundle pricing is very often based on the libraries’ pre-existing print subscriptions, and individual electronic journals are usually priced as a percentage of their print price (typically 90% of print subscription). In addition, many libraries continue subscribing to print copies along with the electronic versions because of VAT taxation differentials. Moreover, while the transition to electronic publishing does represent some specific costs (as well savings, in other dimensions), it is important to stress that this transition cannot be held responsible for the high journal prices that libraries have complained about. While expenses linked to the electronic transition have really started in the mid-1990’s,⁴⁸ data discussed earlier (by Tenopir and King (2000)) show that very significant price increases have taken place much earlier. Kyriallidon et al. (2004, page 11)

indicate that, for the period 1986-2003, the average journal price rose faster (on an annual basis) between 1986 and 1995 than between 1995 and 2003 (see footnote 33). Two conclusions can be drawn from these facts:

1. Over the last 30 years very large price increases are observed, that cannot really be justified by the evolution of costs, since the bulk of the increase happened at a time of pretty stable technology, that is before 1995; this evolution reflects the ability of publishers to take advantage of the relative price insensitivity of buyers, and especially for their most popular journals.
2. In comparison, the last 10 years have been more favorable for consumers. First, prices have risen less fast in percentage terms, even if of course from an already pretty high basis. But more importantly, this has happened at a time where publishers have expended costs to undertake the electronic transition, which has improved access for consumers, as demonstrated by usage statistics (see footnote 34).

Since the price differences identified in this Section indicate significant market power, it is natural to ask how to make the market more competitive and prone to scientific diffusion. The next Sections turn to this question, with special emphasis on pricing policies and on access issues.

⁴⁴ Note that here all variables (except dummies) are transformed into logarithms.

⁴⁵ There may be ground here to go to two simultaneous equations, since citations obtained by a journal may be endogenously determined by the first year of publication of the journal.

⁴⁶ Other than the field dummies, for which an F-test should be run to test for differences, but it is glaring that the differences are significantly large, and the test would reject equality.

⁴⁷ Cornell (1998) finds similar results but he studies only agricultural and biological journals. Moreover he divides NFP into Universities, Societies and Government. White and Creaser (2004) find that Oxford and Cambridge University Presses set lower prices than most FP. Chressanthis and Chressanthis (1994) find that being published by an NFP very significantly reduces the subscription price of an economic journal. Bergstrom and Bergstrom (2004) report that a page published by a NFP is between 3 and 5 times more expensive than a page published by a FP in ecology, economics, atmospheric sciences, mathematics neuroscience and physics. See also Bergstrom (2001).

⁴⁸ As indicated for example by Elsevier in its document submitted in the context of this study. Reed Elsevier, 2005.

4. SCIENTIFIC PUBLISHING IN THE ELECTRONIC ERA: INSIGHTS FROM INDUSTRIAL ECONOMICS

The previous section has documented the significance of price differences across scientific journals. The next sections discuss the evolutions of the market implied by the ‘electronic revolution’. After stressing the key features of this industry which are relevant in this current context, the report turns to pricing policies and access issues.

4.1. TWO-SIDED MARKETS, NETWORK EFFECTS AND BUSINESS MODELS

As stressed earlier, a key feature of the “traditional” reader/library-pay scientific publishing model is its ‘two-sided’ nature and the presence on each side of the same agents, with publishers being intermediaries between authors/producers and libraries, which represent readers/consumers:⁴⁹

- Upstream: authors want to be published and therefore view publishers and journals as *substitutes*, possibly imperfect substitutes when journals carry different reputations.
- Downstream: readers want to have access to the stock and flow of knowledge and therefore view publishers and their journals as *complements* (it is not enough for them to have only a subset of the top journals for example).

Publishers fulfill a double mission for individual researchers:

- *Dissemination* of good research: distribution and production of journals

- *Certification*: in fact provided by upstream and downstream segments: researchers provide services as referees and editors for peer review, and also “vote with their writing” by citing other researchers’ works.

Technological evolution has a different impact on the two missions of academic journals, dissemination and certification. Indeed, while the Internet is not currently affecting the certification process (except for the fact that, with the computation of impact factors and the increased reliance of such indicators in academic incentive schemes, the ‘unavoidability’ of highly-cited journals has been strengthened in recent years), it is making dissemination of existing research much easier (through individual homepages and search engines).

Recent evolutions suggest that certification could be done through other means than journal publishing.⁵⁰ It is an open question at this point whether future business models will continue to bundle these two activities, or whether a separate certification service will emerge, distinctly from dissemination.

The ability of authors to post their articles on web pages or within repositories modifies the role of intermediation. Traditional intermediaries (publishers) are not really needed for publication of these articles (since they are produced by the authors) nor for the diffusion of these articles (which is performed via the www). But even in this ‘open access model’ (hereafter OA model) there is a role for intermediaries to facilitate access to information, to certify it and to ensure that the flow of knowledge that is embodied in the articles posted on the www will be preserved. These activities are costly (see the chapter on preservation of archives in this study) and will be provided by commercial intermediaries only if the return on the activity is positive.

In either model (the traditional reader/library-pay model and the OA model), some intermediation is needed and the sustainability of the business model will reflect the ability to generate network effects, that is, to attract both authors and readers in such a way that both find a benefit of participating. Authors because their reputation benefits from participating and readers because they have access to quality information.

Some results from the literature on *network effects*⁵¹ are worth emphasizing:

- (a) Coordination is required and anticipations play an important role in creating sustainable networks: authors must anticipate that readers will use and cite their work, readers must anticipate that they will have access to good quality information.
- (b) Small differences in market shares or network sizes may generate dominance in the market: a large network has a competitive advantage on a smaller network since the willingness to pay of participants is larger.
- (c) A necessary condition for (b) is that consumers have a cost of shifting from one network to the next.

In the traditional reader/library-pay model, condition (c) is likely to be present. Remember that contracting is made between two intermediaries: publishers and libraries. Libraries have a fixed budget and try to offer to their users the largest portfolio of quality articles consistent with their budget. Given the generalization of contracts whereby a library may commit for up to three years part of its budget for use of the portfolio of a publisher, there are two types of cost of shifting to another network: there is first the penalty imposed by moving away from the contract and then the opportunity cost related to the lack of residual budget given the previous commitments.

In the OA model, condition (c) does not seem to be present *as long as* there is an efficient search engine. It is worth re-emphasizing here that an “efficient” search engine has to be provided. Right now, Google, Msn search, Yahoo and other engines provide this service. From (a), the success of a particular search engine will rest on the beliefs of the users that the ordering of searches corresponds to the information that they are looking for, and that this information is of “quality”. Right now most engines are free for the user. If these engines become indeed the privileged intermediary for OA models, the continuation of their activity may require a shift in their business model, either by relying on advertising or on costly participation, which may in turn increase the costs of shifting for participants.

4.2. NEW TECHNOLOGY AND COMPETITION

The advent of the ‘electronic era’ has a number of consequences.

First, it facilitates access to the flow of knowledge, and entry of new firms for distributing new knowledge. Indeed, authors can more easily bypass traditional “gatekeepers”, thanks to the low cost of establishing websites. A key barrier to entry remains the ability for journals to attract a pool of editors, referees and authors. Even when a new journal can attract a selected group of distinguished academics, it takes time to attract authors since the reputation of the journal is yet to be established.

Second, even if entry of new journals is facilitated, the access to the stock of knowledge is controlled by (historical) owners, partly because they have obtained copyright or exclusive rights from the authors. This poses a problem of access for the provision of services like cross-referencing or cross-access and gives a competitive advantage to large portfolio holders since they can design a platform for cross-referencing without having to contract for access rights. This return to size exacerbates the *natural* network effects that push towards concentration and large portfolios, and that also make barriers to entry an almost inescapable feature of the industry.

We have stressed already the advantage in terms of reputation of existing journals as well as the need for readers to access the entire stock of knowledge. This favors publishers with established and large *portfolios* of journals and could explain the merger activity in the sector. Concentration per se is not inefficient and could in fact be socially efficient; however, it is also an indication of market power as is suggested by the large price increases in the industry.

This effect of concentration is stressed by McCabe (2002) who highlights the race for size, internally or through *mergers*.

McCabe shows that libraries buy journals following a cost-benefit analysis, maximizing the ‘value-for-money’ of the titles they buy. Faced with this library behavior, a publisher will benefit from having a larger portfolio of titles, loses less business when raising the price of one of its titles, since a higher proportion of the libraries that cancel this title will instead buy *another title* from this publisher. Moreover, McCabe has shown that price increases were substantial during the period 1988-2001 and could be partly traced to previous merger activity. Comforting the network effect (b) of the previous section, McCabe also shows that the price increases occurred even if the merger would not have posed problems within the basic merger guidelines since the combined market shares of the merging parties were less than 30 percent.

Potentially, a third consequence of the advent of the electronic era is the risk that incumbents may *endogenously* raise additional barriers to entry, by increasing the switching costs of customers. Particularly relevant here are the changes in publisher pricing policies, i.e., the advent of the ‘Big Deal’ where publishers sell their whole stock at agreed-upon prices for a significant period of time and where getting out of the contract is costly.

These effects are reinforced when the transaction takes place between intermediaries like publishers and libraries and when libraries have a limited budget. In the literature on two sided markets, it is often the case that the profit maximizing strategy for an intermediary is to subsidize access to one side of the market and charge fees to the other side⁵². In the publishing industry, the dominant business model is the subscription model, whereby authors pay nothing or very little to publish their research while most of the publisher’s revenue comes from fees paid by the libraries.

The emergence of electronic access has however also led publishers, both FP and NFP, to experiment with new business models, offering open access journals (financed by charging authors a fee in exchange for allowing the free access to the published article), and creating or contributing to platforms (like SCOPUS,

JSTOR) that facilitate access to a large body of information.

The market is in a transition phase, where the traditional subscription model is still dominant, and where the feasibility and sustainability of new business models is still unclear. In a time where experimentation should be encouraged, it is crucial to evaluate the risk that the natural advantage of incumbency and the reliance on the traditional chain {author → publisher → library → reader} might allow incumbents to leverage their current position in a way that could be detrimental to competition.

We evaluate these issues in detail in the next two sections, focusing first on pricing policies and then on the various dimensions of access to research results.

⁴⁹ It is a tautology that most markets are two-sided. The modern terminology refers to two-sided markets when the participation to the market – access to a platform and prices of transactions – is influenced by the intermediary and when there are significant network externalities.

⁵⁰ For instance, in economics NAJ Economics (NAJ standing for “Not a Journal”) where members of the editorial board provide brief reviews of articles that are freely available on the web; such reviews constitute peer-reviewed publication but authors are not allowed to submit articles. Specialized search engines like Google Scholar provide the number of citations a given article – even posted on an individual web page – is cited by other articles available on the web, whether published in journals or not; such citation numbers could serve potentially as a basis for a researcher’s evaluation.

⁵¹ See for instance, Shy (2001), Shapiro and Varian (1998) and Dewatripont and Legros (2000).

⁵² See for instance Baye and Morgan (2001), Caillaud and Bruno (2003), Ellison and Fudenberg (2003).

⁵³ Submission fees proceed from a different logic since it is a way to force the author to internalize some of the costs it imposes on the editorial board by submitting; a budget constrained author will then submit only if he faces a reasonably high probability of success.

5. BUNDLING STRATEGIES

We have already underlined that the end user wants to have access to a large portfolio of articles and information. Bundling strategies are therefore outcomes of competition between publishers; they are also a source of efficiency in the market. For instance, a journal is in fact a bundle of articles, and has been until now an efficient way to access new results in a field. When there is more than one journal in a field, end users will value having access to all the journals, i.e., would value having access to a bundle of journals, or a portfolio; libraries or portals becoming then an efficient way to satisfy these needs.

With the emergence of digital information, two new products become valuable to the end user: the digital version of articles published in a journal, whether the “first copy” is printed or electronic and also articles that are available on the www but have not gone through the traditional publishing process. The move to electronics has happened in the late 90s and has offered a large array of new and previously unavailable services. For instance, in October 2005, Elsevier's ScienceDirect alone provided a searchable database of 1,800 journals containing over 7 million scientific research articles, abstracts, the full texts and linked references. A citation history is provided for each article (not for all publishers, though). This will probably get even more powerful with CrossRef, whose aim is the linking of several of such databases belonging to competing publishers (like JSTOR, though unlike JSTOR, it will contain the latest articles as well).⁵⁴ Today CrossRef provides an infrastructure for linking citations across publishers and counts 1496 participating publishers and societies. The access to the full version of an article put on these platforms depends on whether the individual or his institution subscribes.

Big deals offer a library access to the whole electronic collection of a publisher in exchange for the library's commitment to a given expense. There are clear advantages to the end user of having access to the whole electronic portfolio, and publishers point out that big deal contracts emerged as a response to the demand of libraries. But there are reasons to fear that the pricing strategy in big deals raises barriers to entry and creates a lock in effect, and the next subsections will discuss this in detail.

We start reviewing the economic literature on bundling. Some of the effects illustrated in this literature apply not only to existing contracts but also to partnerships between different publishers for putting their collections on the same platform. We then turn to the main characteristics of big deal contracts, using published work by Edlin and Rubinfeld (2004), as well as presentations by publishers made at a conference in Barcelona (2005).

5.1. ECONOMIC LITERATURE ON BUNDLING AND DISCRIMINATION

The basic model in the literature allows a firm three possibilities for selling two products A and B:

- sell the goods separately ("standard option");
- bundle the goods and sell the bundle only ("bundle only option");
- offer the goods separately, as well as the bundle ("mixed bundling option").

The existing literature (Adams and Yellen, 1976, Schmalensee, 1984, Mc Afee et al., 1989, Salinger, 1995, Armstrong, 1996, and Nalebuf, 2004) deals with both bundling options and considers both the use of bundling as a price discrimination device (in economics, ‘price discrimination’ is defined as ‘selling the same good or service to different individuals at different prices’) and as an entry deterrent.⁵⁵

Early work has focused on the behavior of a monopoly firm selling different products.

Clearly, monopoly outcomes only give a rough idea of what can happen in a market dominated by a few competing firms which offer individual subscriptions to journals (in printed format only), and their bundle, the so-called "big deal" (which goes together – or not – with the printed version of the journals). The "negative" welfare results that are generated by a monopoly will be less severe in the case of an oligopoly, which better describes the industry structure nowadays. Recent works by Bakos and Brynjolfsson (1999a,b), McCabe (2004), Jeon and Menicucci (2005), and Ginsburgh and Zang (2005a,b) have applied and extended the insights from this literature to the market for scientific publishing to account for competition, limited budget or cooperation between competitors.

Bakos and Brynjolfsson (1999a,b) concentrate on the issue of menus of bundles aimed at different market segments and the incentives to innovate. They point out that traditional price discrimination strategies may be even more profitable if there exists uncertainty in the market. Indeed, selling on different market segments may reduce the role of unpredictable idiosyncratic components of valuations. Bakos and Brynjolfsson (1999b) show that bundlers have high incentives to innovate, while others have low incentives. See also McCabe (2004) and Fay and Mackie-Mason (1999) for related results.

McCabe (2004) is relevant for understanding the effects of the electronic revolution on business models: he shows indeed that the advent of e-journals makes bundling and price discrimination a profit maximizing business model because costs of distribution decrease. McCabe (2004) therefore describes a period of transition from print to electronic and shows how new business models may emerge in response to a decrease in distribution costs.

Jeon and Menicucci (2005) analyze a "future" where only electronic publications are available and the short term consequences of bundling strategies in the presence of fixed budget of libraries. They show that bundling is a profitable and credible equilibrium strategy for firms in the industry.⁵⁶

Moreover, while in the absence of bundling concentration would have little effect on prices, in the presence of bundling, higher concentration brought for instance by mergers will have a positive effect on prices.

Note finally that Ginsburgh and Zang (2005a,b) analyze strategies of creating "super bundles" or "bundles of bundles."⁵⁷ They allow joint ventures between publishers for selling bundles of journals and they compare the three options (standard, bundle only, mixed bundling). While they find that in a small number of situations bundling has a positive effect on consumer surplus, the entry-deterrence dynamics dwarfs this short-term beneficial effect. They also point out that in the case of two firms offering their individual products and a bundle, the various profit sharing rules have different consequences on welfare. They do not consider in detail the long-term welfare effects, which may result from entry-deterrence,⁵⁸ and ignore budget constraints.

Publishers sell their bundles discriminating across subscribers. This discrimination may be based on observable characteristics. See e.g. JSTOR (www.jstor.org/about/us.html) which discriminates between "very large, large, medium, small and very small institutions." The classification is based on objective criteria, that is, the Carnegie Classification of Institutions of Higher Education. Tariffs are published on their website. This is not the case with for-profit publishers who sign confidential agreements with libraries.

5.2. BIG DEALS

As we already pointed out the characteristics of the market and the advent of the electronic era push naturally towards business models where users can access a large set of journals or scientific information. While electronic delivery is becoming more important, standard delivery via printed copies remains integral to libraries' activities. Publishers offer a variety of contracts enabling the libraries to subscribe to both print and electronic versions.

Some publishers (Elsevier, Blackwell, Springer, Oxford University Press) have recently presented the main features of their big deal contracts and their thoughts for future pricing schemes.⁵⁹ Initially, the main contract was “print+electronic”, that is the library committed to continue spending on printed journals what they spent in the past, and the full access to the electronic collection was offered at a surcharge. More recently, both “Electronic+print” and “Electronic only” contracts are offered.

There have been in the US and in Europe some voices to complain about such contracts. In the US, large universities like Harvard, Cornell, Michigan have decided not to sign such contracts and rather pay journals on an individual basis.

Potential worries that big deal contracts raise are as follows:

- (a) If the contract is long term (three-year durations seem typical)⁶⁰ the library has little room to re-adjust its expenses. This reduces the flexibility for libraries to respond to new offerings on the market, which is regrettable in a world with a rapidly changing technology and evolving business models.
- (b) The commitment to spend the historical amount on printed copies makes cancellations less attractive; this is amplified when the savings for canceling a subscription are inferior to the price of an individual subscription.
- (c) Both effects make entry by competitors or new business models more difficult.

The cost of serials and the strong desire of libraries to benefit from full access to the electronic collection has led libraries in Europe (and elsewhere) to organize in consortia, which has also made the pricing of contracts quite complicated, since while all members of a consortium have access to the full electronic collection, they differ in their (historical) purchases of printed journals. An illustration of some of the features of big deal contracts is offered in Edlin and Rubinfeld (2004).

They analyze contracts negotiated with Elsevier in 2002 by individual libraries and consortia in the US. There were three options offered.

- (a) With “Limited collection”, the library chooses for which print journal there will also be electronic access (at a surcharge of 25% of the printing fee).
- (b) With “Complete collection” the library commits to continue purchasing printed journals in the future, the library obtains the access to an electronic version at a 12.5% additional fee for *any title* and benefits from a price protection (a bound on future price increases), swapping of similarly priced titles is allowed, and cancellations are allowed, but at a cost.
- (c) With “ScienceDirect E-Choice,” the library gets full access to all titles for which the library subscribes at a price equal to 90% of the total print subscriptions, print journals are available at 25% of the base print list price.

Options (b) and (c) are examples of big deals, (b) being “Print+Electronic”, while (c) being “Electronic+Print”. A contract like (b) creates an incentive for a library to keep its preexisting *journal portfolio* since it is ensured of a price increase lower than 7.5% per year for the length of the contract while prices for individually purchased subscriptions increase at a faster pace. Contracts (b) and (c) also enable the library to expand its electronic access, either to cover a subject area or the full ScienceDirect database.

As Edlin and Rubinfeld (2004) argue correctly such options make it less costly for libraries with a small historical expense to have full electronic access.⁶¹ Hence big deal contracts offer the possibility to discriminate and allow the publisher to extract more surplus from each library. The literature on third degree price discrimination points out the potential efficiency gains from third degree price discrimination, but it also shows that this discrimination will enable the monopolist to extract most of the surplus from the consumers.

In the case of libraries, this discrimination therefore implies a disappearance of some future options (since the budget is already committed). While an individual library (or consortium) must prefer the big deal contract (by revealed preference) their choice may in fact prevent them from future purchase and therefore may prevent potential entry in the market.

A lock-in effect or switching cost generated by big deal contracts is that while there is a possibility to swap titles, cancellations are made very unattractive since canceling a title brings a saving to the library that is well below the purchase price of the title or that of potential competitors.

Hence, in general, bundling has two welfare effects, (a) in the short term, it may restrict or enlarge choices by consumers, and (b) in the long term it may restrict entry. Even if there are sometimes positive short run welfare effects, the entry-deterrence effect can hardly be avoided once a bundle is in place, especially when as in the current system libraries have fixed and limited budgets and when big deal contracts for some publishers represent a significant proportion of the available budget.⁶²

Table 5 describes a hypothetical situation in which the journals belonging to a domain are sold as a bundle. Each row of the table gives the “share” of a publisher in terms of the price of a paper subscription to its journals. To obtain the “share” of publisher *i* in domain *j*, we divide the sum of the prices of the journals published by *i* in domain *j* by the sum of the prices of all journals in domain *j*. These “shares” are, however, imperfect because some data on prices is missing. Column 2 of Table 1 gives the number of journals per domain and the number of journals for which we have the price (between brackets).

In Biochemistry and molecular biology, for example, one can interpret the “share” in the following sense: Blackwell would make for 4 percent of the price of the bundle, Elsevier would make 44 percent, and all the publishers listed would be in for 77 percent, leaving 23 percent for all others. This would again make entry for new journals quite difficult, since the library budget spent on the domain would probably be exhausted by existing well-cited (ISI) journals.

In all the domains examined, except two (Cell biology and Law) the seven existing publishers appearing in the table make for 50 percent of the bundle. In 10 domains, their share is larger than 75 percent, in 7 additional domains, this share is larger than 60 percent.

⁵⁴ Dirkmaat, J. (2002).

⁵⁵ For instance, Nalebuf (2004) has extended the Adams and Yellen paper by distinguishing the short-term welfare effects due to discrimination and the long-term welfare effects resulting from entry-deterrence.

⁵⁶ They show this even if there is no interdependency among valuations. In their model, when there is no budget constraint bundling would have no effect on prices.

⁵⁷ One could think of CrossRef as a “bundle of bundles” of sort. CrossRef, set up in 2000, is a non-profit association of over 1,400 scholarly and professional publishers that cooperate to provide reference links into and out of their electronic content, and which allows researchers to navigate the online literature at the article level.

⁵⁸ See also Bergstrom and Bergstrom (2004), for similar, though less complete, results.

⁵⁹ Slides made available at ICOLC (<http://www.collectionscanada.ca/consortium/s38-1012-f.html>), Barcelona 2004.

⁶⁰ There are exceptions, e.g., Oxford University Press offers also 1 year contracts.

⁶¹ Their numerical example is as follows. Imagine that the whole print collection in mathematics costs \$100,000. Library A spent \$10,000 while library B spent \$90,000 on printed versions. If the option of getting the full collection is 15% of the print price, library A can have the full collection for a price of $\$10,000 + 15\%(\$90,000) = \$23,500$ while library B pays $\$90,000 + 15\%(\$10,000) = \$91,500$; obviously library B would then have a larger print collection than library A.

⁶² For instance Edlin and Rubinfeld (2004) cite the case of Cornell library which pays 25% of its budget on Elsevier journals (or \$1.5 million) while these represent only 2% of their serials. (“Issues in Scholarly Communication: The Elsevier Subscription,” December 2003, available at www.library.cornell.edu/scholarlycomm/elsevier.html)

TABLE 5 COST OF DOMAIN TO LIBRARIES IF BUNDLED (%)

	Blackwell	Elsevier	J. Wiley	K. Luwer	Lipp.	Springer	T & F	All	Other
1. Biochemistry and molecular biology	4	43	10	9	-	5	6	77	
2. Cell biology	3	14	7	4	-	6	3	37	
3. Chemical, multidisciplinary	ng	12	19	19	-	2	10	62	
4. Chemistry, physical	-	60	8	15	-	3	3	88	
5. Clinical neurology	9	23	4	3	11	20	6	76	
6. Economics	10	51	3	12	-	5	4	85	
7. Education and educational research	11	13	5	7	-	-	32	68	
8. Engineering, chemical	-	54	9	12	-	-	9	84	
9. Engineering, electrical and electronic	-	24	14	7	-	2	8	55	
10. Law	2	11	5	16	-	-	-	34	University Presses
11. Materials sciences, multidisciplinary	-	52	7	11	-	3	4	77	
12. Mathematics, applied	-	39	15	9	-	6	10	80	
13. Mathematics	-	34	6	14	-	15	6	75	
14. Neuroscience	6	40	17	3	3	10	4	83	
15. Pharmacy	3	37	11	2	3	5	8	69	
16. Physics, applied	-	35	2	8	-	8	9	62	
17. Physics, multidisciplinary	-	43	3	11	-	3	3	63	
18. Plant science	12	15	24	2	-	16	7	76	
19. Psychology, clinical	1	19	12	21	-	-	10	63	
20. Psychology, multidisciplinary	3	19	8	5	2	-	18	55	
21. Sociology	12	9	-	12	-	-	17	50	
22. Surgery	4	15	10	-	13	22	4	68	

Notes. Lipp = Lippincott; T & F = Taylor and Francis; All = All those listed (From Blackwell to Taylor and Francis)

5.2.1. CONSORTIA

As stressed above, cost and access considerations have led libraries to organize in consortia. Consortia exist in almost every country. Most libraries regroup on a geographical basis. Their scope is usually linked to the main funding agencies. In Belgium, consortia regroup all universities funded by the French or Flemish Community; in France, COUPERIN, is a national consortium including publicly funded institutions for higher education and research, together with publicly funded research organizations such as CNRS, INSERM and INRA; in Germany, consortia are set up at the Länder-level. Some consortia are subject-oriented (physics, health,...). CSFB (2004) reports that the largest single university buyer of STM journals is the University of California. They represent 2-4% of the global market.

While examining the effect of consortia warrants a specific study, the existing literature suggests the following points.

- One may view consortia as an illustration of “buyer power”, that is the beneficial concentration on one side of the market in order to counter market power on the other side. As the literature shows,⁶³ such cooperative agreements may be socially beneficial, but are also difficult to stabilize when there is heterogeneity among their members (there is a lot of anecdotal evidence of disagreement within consortia leading eventually to individual negotiations; see Gatten and Sanville, 2004 for an analysis of usage across institutions member of OhioLINK).
- Rather than the desire to exert a countervailing power, the motive for creating a consortium may be a simple consequence of the budget constraints of the individual libraries: a consortium pools the budgets of individual libraries and enables them to pay the price required for access to the electronic collection of a publisher. Here we may fear that consortia in fact strengthen the possibility for publishers to charge a high price for their electronic collection.

- The stability of the consortium is linked to the allocation of costs within the consortium, that is whether the final cost will be directly related to observable measures like (historical or current) usage or whether it reflects the bargaining powers of members of the consortium – and how cost sharing within the consortium is resolved – and the bargaining power of the consortium with the publishers.

From a policy point of view, it would be desirable to preserve the short term benefits of these contracts - which enable small and large libraries to have access to a large collection of electronic serials - but also to avoid the lock-in effects of these contracts.

5.2.2. LOCK-IN EFFECTS

An important lock-in effect is related to the dependence of electronic price on (historical) print subscriptions, and the commitment by libraries to a given budget during the duration of the contract. The industry should be encouraged to offer pricing policies that avoid this effect.

Indeed, a key lock in effect is related to the difference between the saving offered when a library cancels a serial from a big deal contract and the actual price of the serial. Pricing policies often put pressure on individual libraries for not canceling their subscription, by offering them savings for *net cancellations* (as opposed to *swaps of journals* with a given publisher) that amount only to a small fraction of the price of the journal.⁶⁴ It seems therefore that two simple recommendations would avoid some of the long term negative effects of big deal contracts on entry and competition are the following.

- The price of the electronic access within consortia or for a single library should not depend on the *historical* number of print subscriptions.

- Prices should be related to the *actual usage* by institutions. If measurement of use is difficult, pricing should be based on variables like number of faculty, students, stock of computers with access to internet, etc. It is very important to stress that overall increased usage over time should not lead to price increases *if publisher costs do not increase as a result of the rise in usage.*

We realize that historical usage by libraries is also a convenient tool for allocating the total cost to the consortium among its members. In practice, it is probable that the actual usage will be correlated with the historical usage for print collections⁶⁵. However, while measurement of usage was rather difficult in the print era, it has become somewhat easier in the electronic era; for instance large publishers like Elsevier collate such data, and there are some recent experiments to homogenize data collection. Assuming that such measures of usage can be developed and used in contracting, the current system can be ameliorated by introducing more flexibility in the contracting terms.

For instance, members of consortia are often heterogeneous in their demands, some wanting a whole collection, others preferring to focus on sub-disciplines; ideally, one would want the first to have a contract for the whole collection and the second to pick and choose its bundle. The current contracts do not allow for enough flexibility in the latter choice. There does not seem to be serious technical constraints preventing a contract “à la carte” that enable libraries select the composition of their bundle of journals titles. Prices could be set for bundles of different dimensions, according to an announced schedule. Little theory exists on this issue, but it clearly relaxes somewhat the constraints imposed by big deals in which consumers have little choice.⁶⁶

⁶⁵ See for instance, Dobson Consulting (1999).

⁶⁴ Edlin and Rubinfeld (2004) cite the example of a subscription to *Journal of Economic Theory* (an Elsevier journal) costing \$2070 per year but saving only 10% or \$207 if a library would cancel the subscription within a big deal. And in the presentation made at ICOCL (2004), representatives from Oxford University Press described their Consortium Premium model as follows: it gives a consortium full access at zero extra cost if all members of the consortium have a print subscription, at a cost of 15% of the full price per institution which does not have an existing subscription and at a cost of 90% per institution canceling its print subscription after 2002. Consider the following (imaginary) numerical example with three libraries in the consortium, and an Oxford journal costing \$300/year for a print subscription. Contrast the following three situations: (i) each library subscribes, the consortium gets free access to the electronic version; (ii) one library does not subscribe at the time of the contract, the consortium has to pay \$45 for electronic access; (iii) the three libraries subscribed at the time of the contract but one library cancels the print subscription, the consortium will then have to pay \$270 for electronic access. Consortia (ii) and (iii) look the same in terms of the number of libraries subscribing to the print version; the difference $$(270 - 45) = 225 can be viewed as a penalty that the consortium must pay for canceling print subscriptions. Alternatively, going from (i) to (iii), the savings for canceling one subscription within the consortium is only \$30 (the bundle print + electronic costs \$900+\$0 to the consortium in (i) but costs \$600+\$270 in (iii)). In this example, if a librarian realizes that the print version is no longer used in his institution and wants to cancel it, the total savings of \$30 is probably very inferior to a subscription to another journal (electronic or not). Hence these penalties for canceling existing subscription reduce the options of librarians to substitute Oxford serials by titles of competitors. In this sense, this type of pricing prevents entry. Note that the same logic applies for such contracts signed with a single library rather than a consortium. This type of problem may be avoided if situations (ii) and (iii) would be treated similarly by the publisher.

⁶⁵ On COUNTER – compliant usage data, see for instance Taylor-Roe and Spencer (2005) and Bevan et al. (2005).

⁶⁶ The only paper that considers this issue, albeit for a different industry (cable television), is by Brown and Alexander (2004). They show that such “à la carte” pricing may have positive welfare effects.

6. VAT ISSUES

The financial burden that high VAT rates bears on libraries' budgets is being denounced by libraries as well as by publishers⁶⁷. It has already been underlined in Section 4 of this report that differences in VAT rates applied to print vs. electronic journals induced a bias in the libraries' decision to continue subscribing to print journals, along with the electronic version. We emphasized several drawbacks of this bias. This Section describes the current situation in Europe and highlights its impact in terms of the libraries' buying power and discrimination across Europe. It also examines possible solutions.

6.1. CURRENT SITUATION IN EUROPE

In most states of the European Union, the VAT rate is lower for printed than electronic publications, the latter being subject to the standard VAT. Annex H of the VAT Directive 77/388 lists the products and services that may enjoy a reduced VAT rate at European Union level. Printed books, newspapers and periodicals, as means of cultural dissemination, are included in the list but the same publications on electronic support are not as they are considered as electronically supplied services (2002/38 EC VAT Directive). Publishers offer both the possibility for libraries to subscribe to an electronic collection only, in this case at a fraction e of the print version price, or to subscribe to a "Print plus electronic" collection, in this case, the additional cost for electronic is only a fraction f of the price of the print version; typical contracts have $e \approx 90\%$ and $f \approx 10\%$. Given this offering, the difference between the two VAT rates induces libraries to opt for the "print + electronic" model (see Section 5). Simple computations⁶⁸ show that even if holding the print collection does not bring social benefits, there is scope for the 'print + electronic option'

to be chosen by the library when the VAT rates are not the same. Moreover, the differential in VAT rates may make it revenue maximizing for the publisher to offer such a choice to the libraries in order to avoid the negative effect of the high VAT rate for electronic collections. Hence, *even if final consumers do not value any longer the print versions* it may be commercially beneficial for publishers to continue to produce printing collections and to distribute them along with their electronic version and it is optimal for libraries to buy print+electronic collections. Under the assumption of no social value for print, such a situation is clearly inefficient (since social resources are wasted to produce and distribute the print version) and is a direct consequence of the arbitrage possibility generated by the difference in the VAT regimes.

Moreover the VAT rates applicable in Europe remain highly disparate and very complex. Not only do the goods and services subject to a reduced rate vary across countries, but the standard and reduced rates are different in the Member States (see Table 6). This discrepancy introduces a distortion in the buying power of research institutions within Europe, as for an equivalent budget some of them will be able to buy more information resources than others depending on their national legislation.

In some countries (such as Sweden and Denmark, for instance), universities and other organisations which do not themselves charge VAT benefit from a tax refund mechanism: once a year, the state refunds the whole VAT to the institutions. This enables them to ignore the VAT surcharge when considering the cost of printed or electronic resources since they know that it will not, in the end, affect them.

Finally, the higher rate applied to electronic delivery of information in Europe strongly affects European research institutions, especially when compared to other countries where electronic services are exempt from tax, such as in the USA. As access to research publications is fundamental to the development of the European Research Area, taxes on scientific information should be reduced to the lowest or even zero rate.

TABLE 6: VAT RATES APPLIED IN EU MEMBER STATES

(Rates are in %)	Printed Books:	Printed Periodicals:	Printed Newspapers:	Standard VAT rate:
Austria	10	10	10	20
Belgium	6	6	0	21
Denmark	25	25	0	25
Finland	8	0/22	0/22	22
France	5,5	2,1	2,1	19,6
Germany	7	7	7	16
Greece	4	4	4	18
Ireland	0	13,5	13,5	21
Italy	4	4	4	20
Luxembourg	3	3	3	15
Netherlands	6	6	6	19
Portugal	5	5	5	19
Spain	4	4	4	16
Sweden	6	6	6	25
UK	0	0	0	0
Slovenia	8,5	8,5	8,5	20
Estonia	5	5	5	18
Slovakia	19	19	19	19
Lithuania	5	5	5	18
Malta	5	5	5	18
Hungary	5	15	15	25
Latvia	5	5	5	18
Cyprus	5	5	5	15
Czec republic	5	5	5	19

6.2. POSSIBLE SOLUTIONS

In order to diminish the VAT impact on research institutions' access to scientific publications and to minimize the discrepancies within and outside Europe, two solutions are examined:

- apply a reduced VAT rate to all type of scientific information, whether print or electronic
- introduce a tax refund mechanism.

In both cases, although the European Commission could play a role in supporting and promoting the solutions, the decisions are ultimately left to the Member States.

The first solution requires that Annex H of the EC VAT Directive be amended. Although the European Commission has presented a proposal for simplifying the rules on reduced rates of VAT in order to ensure its more uniform application, it does not plan to extend Annex H to new categories, i.e. electronic publications. If such an amendment of Annex H was envisaged, to be adopted, the proposal must be endorsed by the Ministers of Finance from the 25 Member States by unanimity.

The second solution, introducing a tax refund mechanism, does not involve a modification of the European VAT rules, and is therefore more easily enforceable, but depends on the willingness of member states to adapt their own legislation. Estimations could be made of the costs that tax refund would represent for the different member states.

As researchers increasingly rely on the provision of electronic publications, the taxes that are levied on electronic information services definitely put a serious strain on the libraries' limited budgets. Reducing these charges would enable research institutions to subscribe, within their limited budget, to more information resources and therefore contribute to accelerate research and the development of the European Research Area.

The tax refund mechanism could be promoted and supported at the European level to urge Member States to adopt this solution for research institutions.

⁵⁷ See for instance the statements of the Frankfurt Group, which includes representatives of authors, publishers & booksellers, libraries, information & research centres, rights organisations and subscription agencies and intermediaries
URL : <http://www.sub.uni-goettingen.de/frankfurtgroup/>

⁵⁸ For instance, if the price of the print version is q , and the VAT rates for electronic and print are $V(E)$ and $V(P)$ then the electronic only collection leads to a final price of $(1+V(E)) \times e \times q$ while the print plus electronic choice leads to a total price of $[(1+V(P)) + (1+V(E)) \times f] \times q$ for the print+electronic collection. Letting $B(E)$ and $B(P)$ be the benefits of having the electronic collection and the print collection respectively, purchasing the print+electronic collection is optimal for libraries whenever $B(P)/q > [(1+V(P)) - (1+V(E)(e-f))]$. Note that even if $B(P)=0$, that is even if holding the print collection does not bring social benefits, there is scope for print+electronic to be chosen whenever $1+V(P) < (1+V(E)(e-f))$; this is never possible when the VAT rates are the same since $e-f$ is less than 1. Note that the revenues of the publisher are $(1+f)q$ in the print+electronic regime while the revenues would be only $e \times (1+f)q$ if the library chose the electronic only collection. Hence the differential in VAT rates may make it revenue maximizing for the publisher to offer such a choice to the libraries in order to avoid the negative effect of the high VAT rate for electronic collections.

7. ACCESS TO RESEARCH OUTPUTS

7.1. ACCESS ISSUES

As the number of scientists grows, so does the amount of scientific articles being published. Information and communication technologies have had a major impact on the dissemination of and the access to research results and their benefits and advantages are being praised by all, scholars, libraries and publishers. Recent research shows that the average number of articles read by university scientists per year has increased from 150 in 1977 up to 216 in 2003; the number of sources used has grown⁶⁹; and enhanced search and browse efficiency enables researchers to spend more time reading the relevant material than searching for it⁷⁰. What researchers are expecting now is to have access to more sources and more backfiles of high quality material; they want speedy access without any barriers⁷¹.

Despite the journal price rises and as a result of bundling strategies and of consortial buying practices (see discussion in previous Sections), the number of electronic journals licensed by libraries has been increasing over the last few years (SQW, 2003; Mabe, 2004). Indeed, electronic consortial licenses have favoured electronic access to all the journals subscribed by the consortium members, or even to the publisher's complete collection, at a comparatively small additional charge.

Usage statistics of electronic journals show that usage (the number of "successful full-text article requests") is rising as well⁷², as a result of the availability of a higher number of journals and of ease of access provided by electronic search facilities and article availability.

Although this entails that "cost per article use" is decreasing, as publishers do stress, effective access to e-journals relies on the libraries' ability to continue paying price-increasing licenses, not only for a growing number of journals (current issues) but also for backfiles; library budgets are under severe pressure, further emphasized by the lock-in effects or switching costs generated by bundles and large portfolios (see previous sections). This raises the question of the long term sustainability of the current scholarly publication system and of the ever-increasing amount of public money necessary to provide access to publicly funded research results.

In view of the libraries' ongoing budgetary difficulties and of the opportunities provided by information technologies, and acknowledging the significant part of public funds involved in the scientific publishing process, awareness of and concerns about scholarly communication issues have been rising in the research community and research funding organizations, and echoed by the civil society and political bodies at national and international levels (Section 1.3 provides a detailed description of the policy context). There is growing concern that in order to maximize research results, they should be widely disseminated and available in open access. Open-access literature is defined as being "digital, online, free of charge, and free of most copyright and licensing restrictions"⁷³. Besides posting them on one's individual web page, there are two primary vehicles for authors to deliver open-access research articles: deposit them in open access archives or repositories, i.e. servers collecting e-prints and/or other scholarly material⁷⁴, and publish in open access journals, i.e. journals performing peer review and making the approved contents freely available to the world. Starting from isolated initiatives, the movement in favour of open access to scientific information has gained scale through the development of open access archives and the growth of open access journals, and the recent policies of research funding bodies about access to research output.

Just like the previous Section was discussing threats and opportunities to make the market more competitive and more prone to scientific diffusion, the remainder of this Section considers how these various dimensions of access can contribute to the same goal:

- Section 7.2 discusses the recent development of open access archives or repositories, their advantages and the questions they raised as an alternative dissemination model.
- Section 7.3 discusses ‘open-access journals’, a more radical alternative to the current reader/library-pay model.
- Section 7.4 discusses the issue of public availability of new research results after publication, especially the possibility for research funders to ‘encourage’ open-access delivery practices.

Section 8 discusses access to the ‘stock of knowledge’, both from the point of view of access to back issues by individual researchers and from the point of view of long-term preservation of research results.

Section 9 addresses the standards and interoperability issues, to ensure data communication and exchange across access systems and platforms.

These various questions are all connected in that they have an influence on the key function of scientific journals in the research process, and therefore on their market power and ability to charge high prices.

7.2. DEVELOPMENT OF OPEN ACCESS ARCHIVES

In the nineties, the development of open source software in universities facilitated the creation of eprint archives and institutional repositories; the best-known programs are Eprints (university of Southampton), DSpace (MIT)⁷⁵.

Two different types of open access e-print archives started to be set up:

- subject-based archives, collecting and providing access to articles and documents in a specific discipline. Their main objective is to allow for a quicker and more efficient dissemination of papers that are deposited by the authors themselves. They have emerged in domains with a long-standing tradition for exchanging pre-prints and where speed of publication is key (physics, computer science).
- Institutional repositories, preserving, disseminating and managing the scientific productions of a specific institution, typically including theses and dissertations, working papers, conference papers and published articles.

In parallel to software developments, the Open Archives Initiative⁷⁶ set up a standard protocol ensuring interoperability between the archives servers. The OAI-PMH (Open Archives Initiative Protocol for Metadata Harvesting), first released in 2001, allows metadata to be retrieved from scattered archives and repositories, and to aggregate this data so that it can be searched with a single query. Services have been set up to help users locate relevant archives, provide unified search facilities and access the documents archived, some of which are discipline-specific, others are generic; for instance, OAIster currently harvests metadata from over 390 archives⁷⁷.

More recently, search engines have also come into play. Scirus, Elsevier’s science-specific search engine on the Internet, which also covers proprietary databases (such as ScienceDirect), started a program to index open access repositories.

Google and Yahoo! have begun to index metadata from various archives and service providers: Google Scholar searches specifically for scholarly literature, from a wide variety of academic publishers, professional societies, preprint repositories and universities, as well as scholarly articles available across the web. Yahoo! started to index data harvested by OAIster in 2004 and has just launched its new Search Subscriptions service enabling users to find subscription content and public web content with a single search.

By enabling researchers to find and access articles available in scattered resources, all these search services help to maximise research access, usage and impact.

In July 2005, the Institution Archives Registry, tracking the number and size of open-access eprint archives, counts 460 archives in the world (of which 222 institutional or departmental, 56 cross-institution, 54 e-theses, 41 e-journals)⁷⁸. The openDOAR Project⁷⁹ is currently building a Directory of Open Access Repositories that will categorise and list the wide variety of open access research archives that have grown up around the world. National approaches are witnessed in various countries: in the Netherlands, the government financed the creation of DARE, the *Digital Academic Repositories* network, providing a single access point to academic research output in the Netherlands⁸⁰. In India, a similar project is planned. In Australia, all the major universities have developed their Institutional Repositories with support of the *Department of Education, Science and Training*. In France, the CCSD, *Centre pour la Communication Scientifique* at CNRS, hosts several open access archives. In the UK, the JISC funded the *Focus on Access to Institutional Resources* (FAIR) Programme⁸¹ under which many repositories have been developed, and has just started the *Digital Repositories Programme*⁸², funding 21 projects, intended to ensure the maximum degree of coordination in the development of digital repositories, in terms of their technical and social (including business) aspects. The i2010 Communication from the European Commission also calls for the creation of digital libraries, focusing on cultural heritage; a Communication foreseen in 2006 will address the accessibility of scientific information⁸³.

7.2.1. MAIN ADVANTAGES OF OPEN ACCESS ARCHIVES

The main advantages of open access (hereafter OA) archives are:

No barriers to access: OA archives provide immediate, free and maximal access to research results, whether published or not, for anyone with an internet connection.

Interoperability of OA archives improves the visibility of their contents, search efficiency and therefore access to research results, retrieval and reading.

In particular, subject-based archives contribute to the fast dissemination of research, pre-prints and post-prints, and provide efficient centralised access to aggregated full-text articles in specific domains.

Institutional repositories contribute to raise the profile of the institutions, making their research output visible and accessible, and provide a potential research assessment tool⁸⁴, for monitoring and evaluating research in institutions as well as at national and European levels. In this respect, the Dutch “Cream of Science”, a subset of the DARE network of repositories providing access to 41.000 publications of over 200 top scientists in the Netherlands, is a showcase of Dutch science and research⁸⁵.

Enhanced visibility and accessibility may lead to higher citations: Although more substantial investigations are needed to confirm this effect, recent studies show that open access increases impact⁸⁶. An analysis at the article level reveals that there is an important difference in terms of the frequency with which the article is cited when it is also available in an OA archives, i.e. when comparing the citation of these articles also available in OA archives with other articles published in the same journals but which are not deposited in OA archives. (Harnad and Brody, 2004).

7.2.2. QUESTIONS RAISED ABOUT OPEN ACCESS ARCHIVES

Disciplinary differences

There are differences between disciplines, both in terms of resources used and means of dissemination of results, which impact the development and use of archives. SPARC mentions 8 disciplines that have successfully set up e-print archives: high-energy physics and mathematics (arXiv), economics (RePEc), cognitive science (CogPrints), astronomy, astrophysics and geophysics (NTRS and ADS), and computer science (NCSTRL)⁸⁷. In chemistry, the preprint server launched by Elsevier did not get the expected success⁸⁸.

In medicine, it is crucial that only fully validated articles be disseminated, and consequently only post-prints are to be archived. PubMed Central (PMC), the U.S. National Institutes of Health (NIH) digital archive of life sciences journal literature, accepts only published articles, and even directly encourages publishers to deposit all their journals contents. Means of dissemination are highly varied across disciplines, be they monographs, patent applications, journal articles, book chapters or conferences. These differences between disciplines of the researchers' behaviour in information search and dissemination have been underlined in a recent survey of 780 UK research academics⁸⁹; they should be acknowledged and taken into account when setting up discipline-based archives and institutional repositories so that they meet the needs and practices of the research community.

Quality of archived documents

According to their purpose, discipline and origin, archives collect different types of research output that has gone through different levels of quality evaluation (working paper, thesis, conference presentation published article...), or none (pre-print). There is some concern (House of Commons 2004, PALS 2004) that the quality validation status of the documents might not be clear to the users; this uncertainty is emphasised when multiple versions of articles are deposited in an archive, and when searching across different archives. Various solutions are already being investigated to indicate clearly the document status (e.g. the House of Commons' report recommends a "kite mark" that can be used to denote articles that have been published in a peer-reviewed journal⁹⁰; JISC is funding the new project "VERSIONS" investigating the issue of versions of Eprints - User Requirements Study and Investigation Of the Need for Standards⁹¹).

Costs

OA archives are set up by researchers, libraries or IT departments in universities and research institutions. The costs of setting up and running such archives are usually being borne by the institutions, sometimes with additional government funding for national or regional initiatives.

Costs estimations are provided for instance in the UK House of Commons' report and in PALS (2004). Although installation costs are low, maintenance costs are more difficult to plan, as they will vary with the number of records and the long-term preservation purposes⁹². Some commercial companies already propose their service for outsourcing the installation and maintenance of repositories for institutions and research organisations⁹³. However the costs remain in charge of the research institutions and organisations, and rely, ultimately, on public funding. In this respect, more work need to be done on the long-term sustainability of open repositories and on the cost/benefit to their founding organisations⁹⁴.

Copyright

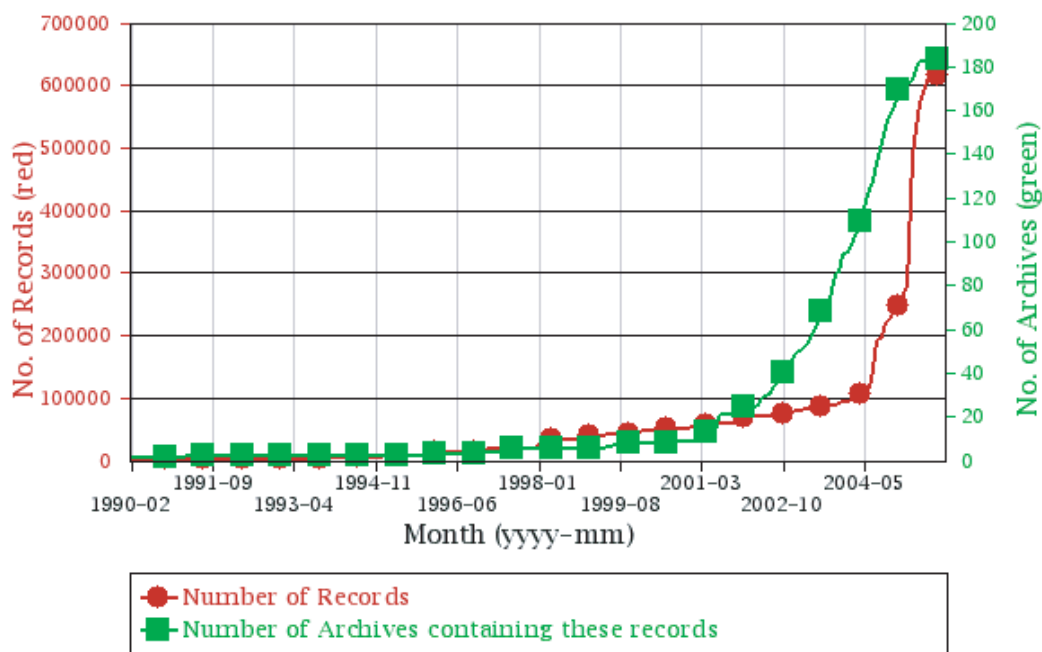
Published articles are submitted to copyright provision as stated in a contract between the author and the publisher, therefore self-archiving must be specifically authorised by the publisher. In July 2005, the JISC-funded SHERPA website registered 116 publishers' copyright policies towards self-archiving, covering 8460 journals; 67% of these publishers allow authors to self-archive post-prints (i.e. final draft post-refereeing), 4% permit archiving of pre-prints only (i.e. pre-refereeing), while 28% do not authorize self-archiving. While publisher policies have become more permissive over time, authors still have to check the publisher's policy for each article to be self-archived. Many variations appear among permissions regarding the version which may be deposited (the exact usage of the terms pre-print and post-print can vary), the type of archive allowed (self-archiving may be allowed in institutional repositories but not in subject-based archives), the delay before public access can be provided to the archived article. Furthermore, the SHERPA database is covering only a proportion of the 2,000 existing publishers (though major ones) and of the estimated 17,700 scholarly journals, so that the picture is actually more complex and currently far from complete.

Quantity

Although the number of OA archives and their contents are globally growing, some of them contain only a small number of documents.

Growth of Institutional Archives and Contents

Generated by <http://archives.eprints.org/>



<http://archives.eprints.org/index.php?action=analysis>

(July 2005)

The success of OA archives relies on the authors' self-archiving of their research papers. As already mentioned in some disciplines pre-print dissemination practices have definitely determined the success of subject-based archives. But the contents of institutional repositories appear to grow more slowly (PALS 2004). Indeed, two recent surveys⁹⁵ show that authors are not knowledgeable about institutional repositories or do not know whether their university has one. However it seems that a majority of researchers would be willing to deposit their articles in an open repository if they were required to do so by their employer or funding body⁹⁶. Some research institutions have started to establish author-archiving policies in order for their institutional repositories to reach the necessary critical mass (see Section 7.4)⁹⁷.

7.2.3. COMMENTS ON CONCERNS EXPRESSED BY PUBLISHERS

Medium-sized commercial publishers seem most reluctant to authorise self-archiving "possibly due to a conscious effort to retain subscription revenue"⁹⁸. Learned societies that make a surplus from their publishing rely on it to support their scholarly activities (conferences, bursaries, grants, public education and society organisation). They fear that open access to articles in subject-based archives might lead to journal cancellations and threaten their existence and activities. The experience in physics has however shown the contrary: the availability in arXiv of articles published by the Institute of Physics and by the American Physical Society has not brought journal cancellations⁹⁹, but may rather have enhanced the journals visibility. As stressed in the UK House of Commons (2004) report, the academic community values several functions of the traditional journal very highly: peer review, subject-specific groupings of articles, access to research conducted around the world. "All this suggests that, in the

immediate term, institutional repositories would not damage the business model on which traditional journals are predicated and that subscriptions would be maintained”¹⁰⁰.

In 2004, institutional repositories did not seem to worry publishers so much¹⁰¹. Some of them believe that open archives can increase their journals visibility. Oxford University Press is currently experimenting open access to 400 articles written by authors from the University and deposited in the institutional archive. They expect their impact to increase. However the recent policies of several research funding bodies in favour of self-archiving raise some publishers associations’ concerns about the potential negative impact on the viability of their journals (see Section 7.4).

7.3. OPEN ACCESS JOURNALS

7.3.1. FACTS

Market

Besides self-archiving, another way for authors to make their publications widely available and freely accessible is to publish in open access journals. Two studies from ISI (Testa and McVeigh, 2004 and McVeigh, 2004) show that open access journals and priced journals have similar citation impacts in their respective fields, but an analysis at the article level seems to indicate that open access increases impact (see section 7.2.1).

The number of open access journals has grown steadily; in November 2005, the Directory of Open Access Journals¹⁰² records over 1900 full text, quality controlled scientific and scholarly journals, in any subject and language, whose current issues are freely available online. These journals can be financed through different sources of income:

- direct or indirect public funding, i.e. by providing subsidies for publishing a journal or by paying the salary of the researchers acting as editors and the infrastructure hosting the journal;

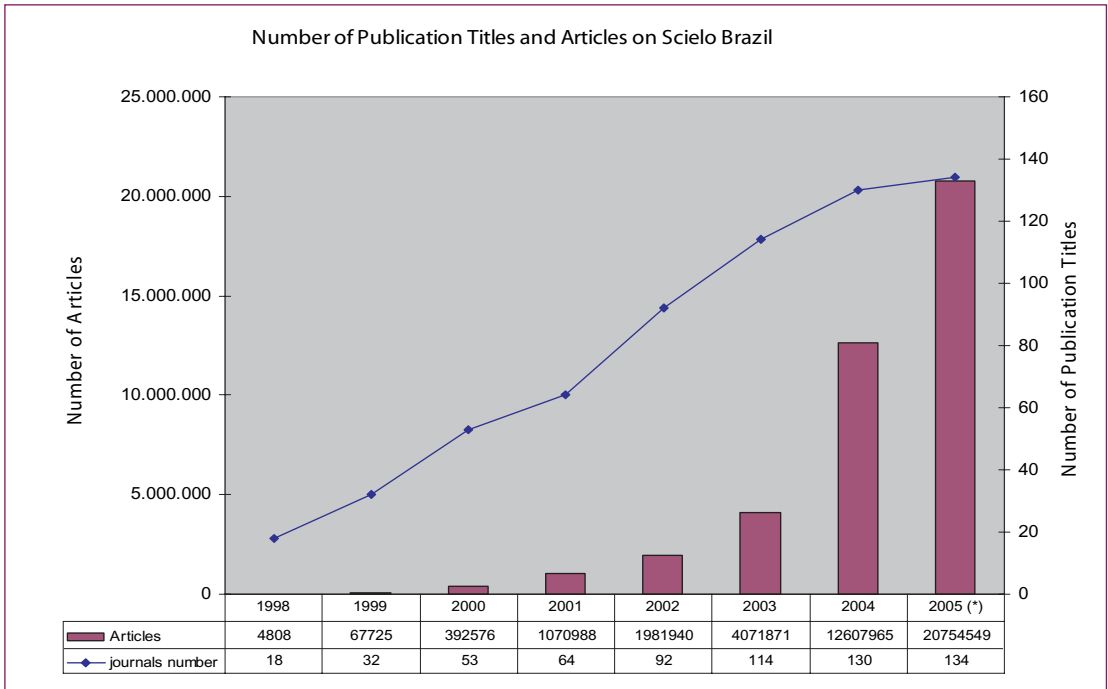
- revenues coming from subscription to the printed version of the journal;
- charging a publication fee (the “author-pays” model);
- and other self-generated income (advertising, sponsorship etc) (Crow & Goldstein, 2003).

In 2004, it was estimated that 55% of open access journals relied on public funding, 28% on print subscription revenues and 17% on the author-pays model (Regazzi, 2004). The subjects covered are mainly health sciences, social sciences, and biology and life sciences (EPIC, 2004).

Public authorities support

In some countries, public authorities have funded large-scale projects to develop portals providing free online access to selected quality scholarly journals published locally. Acknowledging that “scientific journals from developing countries face several distribution and dissemination barriers which limits the access and usage of locally generated scientific information” [SCIELO Web site], some public agencies have launched projects to develop open access electronic journal publishing platforms, in Brazil and other Latin American countries, in Japan and in India, for instance. They aim to provide an efficient way to ensure *universal dissemination* of their journals, and to increase *visibility and accessibility* to their scientific literature. They expect to increase the quality (through a severe selection of titles) and the credibility of regional/local high quality journals, to contribute to the development of scientific research in their countries as well as to maintain and develop science and technology research at an international level. The large scale of the projects is expected to contribute to their success.

SciELO¹⁰³ (Scientific Electronic Library Online) is a particularly successful initiative. Launched in 1997 by several Brazilian public agencies¹⁰⁴, it now provides free access to over 200 selected journals in Spanish and Portuguese, hosted on different national SciELO sites organised in a network. Its operation is highly based on national infrastructures, which contributes to guarantee its future sustainability. In addition to providing access, SciELO provides a model¹⁰⁵ for cooperative electronic publishing of scientific journals on the Internet.



(Source: SciELO <http://www.scielo.org/>, August 2005)

In Japan, the JST (Japan Science and Technology Agency) set up J-STAGE¹⁰⁶ in 1999, a platform for electronic publishing of journals of academic societies sponsored by government. More than 230 journals and 86 proceedings and reports

from Japanese scientific societies in science and technology information¹⁰⁷ are currently published (many of them in English). The journals are still published in print for subscription, but electronic access is provided for free.

J-STAGE TODAY: NUMBER OF PUBLICATION AND ARTICLES



(Source: J-STAGE http://info.jstage.jst.go.jp/eng/info/list/journal_num.html, August 2005)

Public agencies funding such projects believe that open access to research literature achieves a **quick impact** and makes quality article much **more visible and accessible** to an international readership. The higher journal visibility produced by SciELO is likely to have increased the volume of print journal subscriptions [Minon, 2001] as well as their usage, and to have raised their quality profile¹⁰⁸.

In India, the scientific community has launched several open access projects with the active participation of governmental funding agencies, learned societies, and publishers. Several publishers provide open-access to their journals, using government grants and subscriptions to their print version to cover publishing costs (they do not charge authors for publishing their papers)¹⁰⁹.

The BioLine¹¹⁰ initiative is also worth being mentioned: it is a non-profit electronic publishing service for developing countries which helps journals making the transition from print to electronic open access. It already makes freely available peer-reviewed journals from Brazil, Cuba, India, Indonesia, Kenya, South Africa, and Zimbabwe via the Internet.

Hybrid models

Besides fully open access journals, some journal publishers are providing “partial” free access:

- either free access along a “moving wall” principle, i.e. after a determined period of time after publication (a few months or years) (eg. HighWire Press),
- or free access to some articles only. A few publishers are experimenting with a hybrid model for some or all of their journals, allowing authors to make their articles freely available online against payment of a publication fee (Springer Open Choice programme¹¹¹, and Blackwell Online Open service¹¹²).

Note that STM publishers have also formed and been significant contributors to projects such as HINARI and AGORA to ensure broad access to core health and agriculture materials for developing countries¹¹³.

Impact on European research ?

Different cultural approaches to science communication will have a global impact. *“At present, publishers located in the Asia Pacific are producing the largest number of OA journals selected for coverage by Thomson ISI. Nearly 15% of the covered titles from Asia-Pacific are available as OA, and over 40% of the titles from Central or South America are OA journals”* (ISI, 2004b). It is expected that national approaches to open access journals be adopted in the near future in countries like India and China (Guédon 2004).

As an increasing volume of high-quality research output from outside Europe becomes openly accessible, it raises the question of the visibility and accessibility, and of the subsequent potential impact, of European research, as most articles by European researchers are published in subscription-based journals.

Although national and regional open access projects appear in some European countries, they focus on the provision of open access e-print repositories (France, Italy, the Netherlands...) rather than on the large-scale provision of open access journals. These developments must be considered in the light of the specific structure and organisation of the scholarly publishing markets in the various countries; Minon and Chartron (2005) show that, regarding the SSH publishing markets, their organisation and distribution among private and public actors vary significantly between countries (France, Italy and Spain). The fact that the major European STM publishers are commercial companies is not irrelevant to the developments observed in Europe, i.e. of e-print archives rather than of national platforms with free access to journals.

From a policy point of view, it is important for Europe to make sure that there is a ‘level-playing field’ in terms of business models for scientific journals, in order to allow for an optimal degree of dissemination, and influence, of European research.

7.3.2. COMPETITION IN THE READER/LIBRARY-PAY AND AUTHOR-PAY SYSTEMS

Within the existing reader/library-pay system of diffusion, libraries play the role of “budget minister” for universities, institutions and researchers and negotiate directly with publishers. Given this contracting mode, competition for the library revenues is made in a market where journals are viewed as complement rather than substitute: as we have shown, this magnifies the benefits of portfolio size and also makes competition rather ineffective in decreasing prices. This lack of price discipline is exacerbated by the passivity of the research community, both in communicating to the librarians their needs and in not internalizing the cost of accessing publications.

Beyond making research results accessible through individual web pages or repositories, there are two (non conflicting) theoretical possibilities for increasing price competition in the market: shift price competition to a level where journals are viewed as substitute rather than complement or make researchers and users more price sensitive. Author-pay systems or other reader-pay systems like pay-per-download may achieve these goals.

In an author-pay system, where authors pay for being published in exchange for a free access of readers to their article, competition shifts to the market for authors. Everything else being equal, authors should prefer to submit and publish in journals charging a lower price. In a pay-per-download system the reader may also be more price sensitive but the price effect there is less clear since journals are viewed as complements, and since the equilibrium price per download is likely to be small. There are instances of such offerings on the market and JISC is currently experimenting with hybrid systems combining both author-pay and pay-per-download systems.

There are benefits and costs associated to reader-pay (the current subscription system or pay-per-download), or author-pay systems. We gather in the following table some of these trade-offs. The magnitude of these effects is difficult to quantify at this point because they are highly dependent on the way the industry

actors will coordinate. For instance, both pay-per-download and author-pay systems require that the researcher has access to a budget (sometimes significant since the current price for an author-pay article is around 3000 Euros.) This requires political intervention by either injecting extra budget towards publishing and access to researchers or re-allocating some of the existing budget between libraries and researchers.

Recent data discussed in the previous subsection suggest that articles published in open access journals are more cited than articles in other journals. This result is consistent with the view that a zero access cost for readers should increase diffusion. It is well known that “network effects” can be more easily created when one side of the market is heavily subsidized (e.g. Baye and Morgan, 2001). This is true for the current reader-pay system by which authors do not pay for the certification service provided by the publishers; however, in this case, network effects benefit publishers who can attract a lot of authors, that is who have a large portfolio of journals. In author-pay, all researchers are potential readers and this network effect benefits all journals and all publishers, sometimes independently of their portfolio. This suggests that author-pay may facilitate entry in the market, by limiting the competitive advantage given to publishers with large portfolios.

The recent literature on pricing on two sided markets (e.g., Rochet and Tirole, 2004 for an overview) points out that it is important to consider both the platforms’ fixed charge and the usage fee. In the current system, usage pricing is done in a roundabout way since the subscription of a library to a journal or the big deal contracts contains an *ex-ante* evaluation of the future usage of the journal. By contrast, a pay per download system will be closer to an *ex-post* usage. Hybrid systems where an institution pays a fixed fee for the right to access a collection and final users pay a per-download fee for actually getting access to a particular article are made possible by the new technologies. The theoretical implications of such hybrid systems are unclear as are the benefits of heavily subsidizing one side of the market.

One advantage of a hybrid system over author-pay system is the ability to cover costs. Indeed, the revenue base of an author-pay system is by nature smaller than the revenue base of a reader-pay system; since an author pay system strengthens price competition (for authors), and since electronic publishing involves high fixed costs and low marginal costs, it is a natural consequence that author-pay systems will find it more difficult to cover costs than reader-pay or hybrid systems.

Another advantage of an hybrid system is that both the supply and the demand for articles are price sensitive, and therefore the emergence of new offerings on the market may be reflect better their social value. For instance, a new journal offering frontier research in biology may not be economically sustainable in a pure author-pay system while its social value for the biotech industry and society may be important. Having the industry pay part of the cost of publication, via a pay-per-download or a subscription system, would therefore enhance cost recovery and enable socially beneficial entry. When we point out in the table that a negative aspect of author-pay system creates a subsidy to the industry, we want to highlight both the fact that some public funds could be freed (since authors may have to pay less) but also that the entry of new journals may be facilitated when such a subsidy disappears.

These trade-offs are qualitative and our main conclusion at this point is to stress the need for further study in order to quantify these costs and benefits and to assess the feasibility and desirability of these alternative models (note moreover that in the future they may very well coexist across various segments of the publication market).

The industry is fast changing and from discussions with the actors it seems clear that the future is still uncertain. Most of the economic analysis in the first part of this report has been within the traditional distribution system, based on subscription and negotiations between publishers and libraries. As we have pointed out in the previous section, the prevalence of this system makes it difficult to evaluate the feasibility of alternative systems, principally because such feasibility depends on how the overall budget for publication and diffusion is allocated. It is worth noting that, if the research funding authorities want to ‘give a chance’ to the author-pay model, they have to allow for a ‘level-playing field’ in comparison with the reader/library-pay model, that is, provide funding for publication costs and not only for library budgets.

SOME PROS AND CONS OF DIFFERENT BUSINESS MODELS.

Pay-per-Download		Author Pay	
<i>Publishers viewed as complement</i>		<i>Publishers viewed as (partial) substitute</i>	
Pros	Cons	Pros	Cons
Price sensitivity of the reader	Financing (ability to pay of the reader)	Price sensitivity of the author.	Financing (ability to pay of the author)
Metering of usage and willingness to pay of readers	Implementation of the system and security of payment	Better dissemination (readers have free access)	May create wrong incentives for quality publishing (cater “too much” to authors) ¹¹⁴
Competition shifts to the article level (limits bundling effects).	In the short term, more risk on publishers (in terms of end demand)	More price competition	Indirect subsidy to industry

7.4. RESEARCH FUNDING BODIES' POLICIES CHANGING CONDITIONS OF ACCESS

Following the declarations in favour of open access to publicly funded research results, the development of open access archives and the emergence of new business models, several research funding bodies have established or announced policies supporting open access to the research output they have funded. They call upon their researchers to publish in open access journals and to deposit the articles arising from the funded research in an open access repository:

- The US National Institutes of Health (NIH) request and strongly encourage their funded researchers to submit their articles to PubMed Central¹¹⁵ as soon as possible after publication, and no later than 12 months after (policy effective since May 2, 2005)¹¹⁶.
- The UK Research Councils (RCUK), the public agencies funding research in all disciplines, issued its “Position Statement on Access to Research Outputs” in June 2005. The draft policy calls for mandating the RCUK grantees to deposit the results of their research, published articles and conference presentations, in open repositories at the earliest opportunity. Research Councils UK has undertaken extensive consultation with key stakeholders and aims to be able to publish an updated position statement in early 2006¹¹⁷.
- Even a private organisation like the Wellcome Trust, the UK's biggest non-governmental funder of life sciences research, mandates its grantees to deposit their articles within 6 months of publication, either with PubMed Central (PMC in what follows) or a future UK/European version of PubMed Central to be established (policy effective as of October 1, 2005)¹¹⁸.

Other important research-funding agencies in Europe have announced and/or implemented similar incentive policies such as CERN¹¹⁹, CNRS in France¹²⁰, DFG¹²¹ and Max Planck Society¹²² in Germany.

Most of these research institutions also encourage their researchers to publish their articles in open access journals where a suitable journal exists, and offer to pay the publication fees for those charging such fees (the “author-pays” model).

By doing so all these research funding agencies expect to accelerate research on the questions they are funding, increase the impact of their research, and the return on their investment in that research. For the Wellcome Trust, having all its research in PubMed Central (hereafter PMC) will “improve the efficiency of strategy setting -for example, setting funding priorities- assessing the outputs of the funded research, and even gaining an insight into the impact of the work”¹²³. For public agencies, such policies will ensure that publicly-funded research output is openly and widely available in order to maximise the impact of their investment in maintaining and improving the research base and to increase the contribution it makes to the benefit of the society and the economy¹²⁴.

The different policies show variations on the author's obligation (*request vs. mandate*) and on the length of the maximum delay after publication upon which the self-archived articles have to be openly accessible. Clearly, the NIH's initial “mandate” was softened into a “request” and the maximum delay was extended from 6 to 12 months after considerable discussion with scientific societies and other publishers. The RCUK draft policy received more than 70 responses, some supporting and others opposing the proposal. Publishers fear that as articles become freely available in open archives and as search, access and retrieval facilities are enhanced by search engines and interoperability, journal subscriptions will be cancelled, thereby undermining the viability of their journals. This might, in turn, decrease the profitability of commercial publishers and reduce the surplus available for scientific societies (which they use to cross-subsidize other activities)¹²⁵. Several journals and publishers announced their policy on NIH-funded authors, most of them authorising self-archiving of post-prints but imposing a six or 12 month embargo before authorising public access to the articles (i.e. turning the maximum delay estimated by funding agencies into an embargo).

In some cases, publishers offered to deposit post-prints in PMC on behalf of authors. In response to the RCUK statement not specifying any deadline, ALPSP “recommend to respect the wish of some publishers to impose an embargo of up to a year (or, in exceptional cases, even longer)”, underlining that “no one embargo period fits all – a journal in a fast-moving field, and/or with frequently published issues, may be less threatened than one in a slower-moving or smaller field with, for example, quarterly or bimonthly issues” (Morris and Powell, 2005).

From the researchers’ point of view, although recent surveys show that their self-archiving practices are unequally spread among disciplines, a majority of them in all disciplines favour research funding bodies mandating self-archiving (Sparks, 2005). This does not entail that they will turn away from journals; it seems clear that as long as authors and readers value the rich online functionality added by publishers they will still wish to have access to the journal and pay for the subscription. They also believe peer review of crucial importance (Rowlands and Nicholas, 2005) and consider the journal article as the dominant means of dissemination in terms of influence on the Research Assessment Exercise (together with monographs in arts and humanities, language and area studies) (Sparks, 2005). Self-archiving provides them with an additional means of dissemination likely to expand access to their research findings, their usage and impact.

With their open access policies, these funding agencies are raising the researchers’ awareness about their responsibility for making their findings widely available and accessible, as well as about the economics of the publishing market (e.g. when comparing publication fees of journals). It can be expected that such policies will contribute to work towards more balanced copyright contracts and, indirectly, to introduce competition between publishers regarding the social value they provide to the research community. As exemplified by the impact of researchers’ assessment on the means of dissemination these researchers aim at, research funding agencies have a central role in determining the researchers’ publishing practice and attitudes to alternative forms of publication.

7.5. CONCLUSIONS

As illustrated in this Section and underlined in the OECD report 2005, ICTs offer opportunities to develop new scholarly communication systems that serve researchers, research users and research funders more effectively, and which increase returns on R&D investment and enhance innovation.

Self-archiving and open access repositories

Among the opportunities, there is the possibility for authors to self-archive their publications in open repositories enabling free and wide dissemination and access of research results by researchers and the public. Indeed, open archives or repositories provide complementary scholarly dissemination and communication services; they should be designed so as to adapt to the needs of researchers, research institutions and funding bodies, and to meet research practices in the different disciplines (e.g. fast communication of pre-prints; central access to full-text articles, conferences, thesis; central access to peer-reviewed articles only, etc.). They will improve the visibility, accessibility and impact of research outputs at institutional, national, European and worldwide levels and provide potential tools for research monitoring and policy making. They might contribute to balance the national approaches by non European countries supporting large-scale open access journals publishing.

In this context, there is a central role for research funding bodies in Europe in the shaping of new models for publishing and communicating research results. They could stimulate the debate about access to scientific publications (as stressed in a recent DFG survey¹²⁶) and define policies that ensure access to research outputs.

Funders can contribute by requiring from their funded researchers that they deposit their published articles in open archives (i.e. self-archiving) and/or publish their articles in open access journals where a suitable journal exist. Regarding self-archiving, for such policies to be effective in all domains, it would gain to be

mandated (as a condition of grants), to ensure that articles are archived in all domains, and not only in domains where authors have a long-standing tradition of exchanging e-prints or are sensitive to the wide dissemination and visibility of their articles.

In order to maximise dissemination and access, self-archiving could be made shortly after publication. Depending on domains and type of journal, a minimum deposit delay might be discussed with publishers, so far agreed to be of up to a year (but allowing for exceptions), before self-archived papers should be made publicly accessible (Morris and Powell, 2005). Though this measure will hinder the direct impact of research results, it is expected to prevent jeopardizing in the short run the scholarly publication system as it currently functions (i.e. by undermining sales, the publishers' source of income).

The NIH experience shows that publishers have, in some cases, extended their previously determined embargo, but in most cases the policy led publishers to authorise self-archiving of post-prints where only pre-prints were previously accepted (Suber 2005). Notwithstanding the usefulness of the SHERPA database, the current situation regarding self-archiving authorisation is complex for users, as publishers show many variations on what they exactly authorise (see above regarding pre vs. post-print, subject vs. institutional repositories, length of embargo); such policies are likely to harmonise the situation and to provide an incentive for researchers to self-archive their articles.

There is also a role to play for *publishers which* could deposit themselves articles in open repositories, benefiting thereby from long-term preservation guarantees where they exist (e.g. PubMed Central). Such an “open archive article deposit” service, which could be offered by publishers to the authors, might be valued by the research community in the “rating” of journals and of publishers as contributing to the social welfare.

Thanks to the “imperfect” nature of the market and the natural barriers to entry, publishers should not be threatened by this type of funders’

access policies; open repositories will not contain the complete research output and will not provide as coherent and professionally-organised information as journals do, at least not in the short term. However market competition should be improved in order to minimise the risk of such policies for small to medium publishers. Large publishers will be able to set up advanced electronic added-value services, on the basis of their high-performance e-journal platforms, to continue to attract users and maintain their revenue, which might not be the case for smaller publishers. To be specific, in some cases, public support might help national or regional quality journals (e.g. in Social Sciences and Humanities and/or for non-English-language journals) make the transition from print to electronic, and become competitive on the electronic market. Such initiatives require further investigation of the structure and organisation of the publishing market for journals in SSH which are quite different between countries, especially the distribution among private and public actors (see for instance the study by Minon and Chartron (2005) comparing the SSH journals market in France, Italy and Spain).

To sum up, specific actions at *the European level* can contribute to improve visibility, access to and impact of European research (whether funded at EU, national or regional levels):

- Establish a European policy mandating published articles funded from the European sources to be available in open access archives, for instance by means of authors’ self-archiving;
- Explore with Member States and with European research and academic associations (e.g. EUA, ESF, EUROHORCS, Allea, Academia Europaea...) on whether and how such policies and open repositories could be implemented by research funding bodies and research institutions in the Member States;
- Specify the standards under which these open access archives must be accessible to ensure interoperability and cross-searching facilities;
- Set up a general European archive for

researchers who do not have access to subject-based archives in their discipline or to an institutional repository.

At the *worldwide level*, the recent CODATA's initiative "Global Information Commons for Science" proposes a partnership and a platform to coordinate action, "for the purpose of raising awareness on the part of the actors and increasing the effectiveness of the activities directed to facilitating various methods of open access and re-use of publicly-funded scientific data and information, and to promoting cooperative sharing of research tools and materials among researchers"²⁷. The undertaking already has significant institutional support from ICSTI, ICSU, INASP, TWAS, UNESCO, and the OECD, and is expected to be launched as an outcome of the second and final phase of WSIS.

New journal business models and rankings

In terms of the future of the industry, note that possibilities opened by electronic access mean that in journal publishing a variety of flexible business models can be tested and promoted. If research funding agencies call for publication in open access journals, they should provide funding for publication costs. This would allow for a level-playing field for the different journal business models to be tested and assessed regarding their economic viability as well as other dimensions. The specific added value of journals typically relies now on peer review management, bundling of articles relevant to a particular community, "branding" level of quality, as well as rich search and linking functionalities. Thanks to the increased availability of statistics, further competition between journals could now be introduced on the authors' side regarding what they value (or what their funding bodies value) beyond citation potential, like speed of publication, self-archiving allowed upon publication or within short delay, author-friendly copyright contracts, abstract and indexing services, reference linking, long-term preservation provisions etc. Given the increased importance of journal rankings on researcher careers, it would be desirable to think about more comprehensive rankings that would raise author awareness of these dimensions by taking them explicitly into account.

Scientific quality, approximated for example by citation counts or impact factor, should obviously remain the predominant dimension of journal quality for scientific evaluation, but it would be useful to design *rankings of 'social responsibility' of scientific journals* in terms of their contribution to good practices of efficient dissemination and preservation. There is a central role for funding bodies to define policies towards their funded researchers which will improve access and dissemination of publications, especially in terms of self-archiving requirement, copyright contract model, "social responsibility" rankings of journals. There could be an impetus from public authorities at the European level for such an initiative, which would add diversity to the existing information supplied by Thomson ISI, and would naturally induce publishers to stress good practices in these dimensions.

Unbundling certification and dissemination

Let us end with a more speculative note on the effect of electronic access concerning the long-term future of the traditional model where publishers fulfill both the roles of certification and of diffusion. There are at least two ways by which certification and diffusion may be unbundled: repositories and browsing services.

Repositories can be used potentially to unbundle diffusion from certification. Repositories can be used to deposit published articles, as it is the case for instance with PubMed Central where papers having benefited from NIH funding will be deposited one year after publication. In this case, certification is provided by the publisher before the article is deposited. But repositories are also used to archive preprints or manuscripts; and here there is no explicit certification, except perhaps (an imperfect) one obtained by measuring the number of downloads of a given paper. While it is possible to envision a certification made on the basis of the archived paper (some recent experiences are NJE in economics and Faculty of 1000 in Biomed), it is not clear yet how such unbundling would operate if based on a peer-review system.

Traditional publishers have developed platforms where the user can browse through large collections in which articles are cross-linked¹²⁸. In parallel, the availability of e-prints, that is, articles in electronic form allows other entrepreneurs to offer such a browsing service; the most well known entrant is Google Scholar. The quality of the service provided to the users by such browsing services depends on whether the user will eventually get the whole article. This in turn is a function of whether authors archive their manuscripts in repositories, of whether research funding bodies force public funded research to be freely accessible after a short period, of whether the author has relinquished his copyright to the publisher. Browsing services can therefore provide another way to distribute certified (or not certified) unbundled articles; as such they increase total welfare.

For each of these unbundling strategies, their feasibility requires that the certifier in the case of repositories or the new distributor in the case of browsing services can get a positive return on investment. At this point it is not clear whether the business models that could generate such positive returns (e.g., paid advertising for browsing services) would not create their own distortions. We prefer to point out the need for further study in this area.

⁶⁹ Tenopir (2005).

⁷⁰ Reed Elsevier Submission to DG Research study on scientific publishing markets in Europe. January 2005

⁷¹ Tenopir (2005).

⁷² See for instance UK House of Commons, 2004.

⁷³ Peter Suber, A Very Brief Introduction to Open Access. URL: <http://www.earlham.edu/~peters/fos/brief.htm>

⁷⁴ In the EPIC (2004) report, a distinction is drawn between “e-print archives”, which refers to a collection of material published in journals, and “e-print repositories” which defines a collection of material including grey literature and published material. However, in the literature both terms are being used as generic terms as well, referring to any server of e-prints and/or other material, whether published or not. This is also the case in this report, unless further specified.

⁷⁵ EPIC (2004), PALS (2004)

⁷⁶ <http://www.openarchives.org/>

⁷⁷ See also ARC Citebase Search and complete list of service providers at <http://www.openarchives.org/service/listproviders.html>

⁷⁸ <http://archives.eprints.org/index.php>

⁷⁹ This project is supported by JISC, the Open Society Institute, SPARC Europe and CURL. URL: <http://www.openoar.org/>

⁸⁰ As of May 2005 DAREnet harvests over 40.000 digital records from the Institutional repositories of sixteen institutes. <http://www.darenet.nl/>

⁸¹ http://www.jisc.ac.uk/index.cfm?name=programme_fair

⁸² http://www.jisc.ac.uk/index.cfm?name=programme_digital_repositories

- ⁸³ EC Information Society. i2010: Digital Libraries, 2005.
- ⁸⁴ See JISC funded project Institutional Repositories and Research Assessment investigating and developing institutional repository infrastructure for EPrints and DSpace to enable Research Assessment, specifically for the UK Research Assessment Exercise 2008.
- ⁸⁵ About 60% (25,000) is full content available. Unfortunately the rest has copyright restrictions.
- ⁸⁶ <http://opcit.eprints.org/oacitation-biblio.html>
- ⁸⁷ Crow (2002) ; CUL (2004)
- ⁸⁸ <http://www.sciencedirect.com/preprintarchive>
- ⁸⁹ Sparks (2005)
- ⁹⁰ House of Commons, 2004. Section 135.
- ⁹¹ JISC funded project VERSIONS
http://www.jisc.ac.uk/index.cfm?name=project_versions
- ⁹² See example of arXiv, transferred from Los Alamos to Cornell University because of running costs, even though they were very low.
- ⁹³ E.g. Open Repository, the service by BioMed Central; Digital Commons@ by ProQuest.
- ⁹⁴ Heery and Anderson (2005).
- ⁹⁵ Sparks (2005), Rowlands and Nicholas (2005)
- ⁹⁶ Swan and Brown (2005), Sparks (2005)
- ⁹⁷ See full list at <http://www.eprints.org/signup/fulllist.php>
- ⁹⁸ Cox & Cox (2003)
- ⁹⁹ Swan (2005).
- ¹⁰⁰ House of Commons (2004), § 139, 260, 261. Publishers have the same worries for abstracts and indexes, which are some times viewed as substitutes of the original paper (see John Cox Associates Ltd, 2004).
- ¹⁰¹ PALS (2004), House of Commons (2004), CSFB (2004)
- ¹⁰² <http://www.doaj.org/home>
- ¹⁰³ <http://www.scielo.org/>
- ¹⁰⁴ SciELO developed by the Fundação de Amparo à Pesquisa do Estado de São Paulo, in partnership with BIREME - the Latin American and Caribbean Center on Health Sciences Information. Since 2002, the Project is also supported by CNPq - Conselho Nacional de Desenvolvimento Científico e Tecnológico as well as national and international institutions related to scientific communication.
http://www.scielo.org/comodel_en.htm
- ¹⁰⁵ <http://www.jstage.jst.go.jp>
- ¹⁰⁷ Fields of biology, chemistry, engineering, medicine, physics, science, and zoology
- ¹⁰⁸ Testa and McVeigh (2004) study on OA journals stresses the quality of open access journals available on these platforms. "The journals of the SciELO project, for example, are among the most venerable OA journals, having been freely available online since 1999."
- ¹⁰⁹ All 10 journals of the Indian Academy of Sciences (URL: <http://www.ias.ac.in>), for example, as well as the four journals of INSA (Indian National Science Academy URL: <http://www.insaindia.org>) are OA journals. The 22 biomedical e-journals of the Indian Medlars Centre of the National Informatics Centre are also accessible without subscription. An initiative by the National Informatics Center (NIC) and Indian Council Of Medical Research (ICMR), two governmental agencies. URL <http://indmed.nic.in/>
- ¹¹⁰ BioLine international: (<http://www.bioline.org.br>) This is a collaborative initiative of scientists and librarians of the University of Toronto Libraries, Canada, Brazil, and Bioline, UK. It is a non-profit electronic publishing service committed to providing open access to quality bioscience research published in developing countries. It makes available published information from peer-reviewed journals from Brazil, Cuba, India, Indonesia, Kenya, South Africa, and Zimbabwe via the Internet. By 2005, the site will host 15 medical and bioscience journals digitized by Medknow Publications, Mumbai. Medknow Publications, a small company based in Mumbai, has helped 10 medical journals — including the Journal of Post Graduate Medicine and Neurology India — make the transition from print to electronic open access and all of them are doing much better now than before.
- ¹¹¹ <http://www.springeronline.com/sgw/cda/frontpage/0,11855,5-40359-0-0-0,00.html>
- ¹¹² <http://www.blackwellpublishing.com/static/onlineopen.asp?site=1>
- ¹¹³ The HINARI program, Health InterNetwork Access to Research Initiative, set up by WHO together with major publishers, enables developing countries to gain access to one of the world's largest collections of biomedical and health literature. URL: <http://www.who.int/hinari/en/>
The AGORA site, Access to Global Online Research in Agriculture, provides access to 800 journals from major scientific publishers in the fields of food, agriculture, environmental science and related social sciences. AGORA is available to students and researchers in qualifying not-for-profit institutions in eligible developing countries.
URL: <http://www.aginternetwork.org/en/>
- ¹¹⁴ See McCabe and Snyder (2004) however on the ability of open-access journals to commit to maintain quality by charging submission fees and not solely publication fees.
- ¹¹⁵ PubMed Central is the U.S. National Institutes of Health free digital archive of biomedical and life sciences journal literature. URL: <http://www.pubmedcentral.nih.gov/>
- ¹¹⁶ Public Access NIH Policy <http://www.nih.gov/about/publicaccess/index.htm>
- ¹¹⁷ Over 70 responses to the proposal have been received. Access to Research Outputs, RCUK Open Access Position Announcement and information on the consultation URL <http://www.rcuk.ac.uk/access/index.asp>
- ¹¹⁸ Wellcome Trust position statement in support of open and unrestricted access to published research
http://www.wellcome.ac.uk/doc_WTD002766.html
- ¹¹⁹ 'Continuing CERN action on Open Access' (<http://doc.cern.ch/archive/electronic/cern/preprints/open/open-2005-006.pdf>), released by the Scientific Information Policy Board (SIPB)
- ¹²⁰ CNRS <http://www.eprints.org/events/berlin3/ppts/10-LaurentRomary.ppt>
- ¹²¹ DFG Status report <http://www.zim.mpg.de/openaccess-cern/presentation-fourmier2.pdf> and <http://www.eprints.org/events/berlin3/ppts/05-JohannesFournier.ppt>
- ¹²² Open Access Policy of Max Planck Society
<http://www.oai.unizh.ch/symposium/docs/Botz.pdf>
- ¹²³ Terry (2005)
- ¹²⁴ RCUK policy #3
- ¹²⁵ See STM reaction to NIH policy STM Position on NIH Open Access Proposal 12 November 2004 <http://www.stm-assoc.org/statements/accessprop.php> and ALPSP response to the UKRC position statement <http://www.alpso.org/events/2005/RCUKmeeting/RCUKResponse.pdf>.
- ¹²⁶ Deutsche Forschungsgemeinschaft (2005).
- ¹²⁷ CODATA "International Council for Science : Committee on Data for Science and Technology", "The Global Information Commons for Science Initiative". Press Release, 1 September 2005. <http://www.codata.org/wsis/Global-CommonsforScienceSept1.html>
- ¹²⁸ Science Direct is an example; recently North-Holland has developed a web interface for its Handbook in economics where the user can click on the articles referenced in the bibliography: since handbooks represent the state of the art in an economic discipline, the value of such an interface is obvious.

8. ELECTRONIC JOURNALS: PERMANENT ACCESS AND PRESERVATION

The electronic era has brought a major paradigmatic change in the provision of access to back issues of journals: in the print era, libraries were *acquiring print journals and took in charge their preservation* so that they remain accessible to their user community in the long term. In the digital era, libraries and their user community are *licensed online* access to electronic journals for a determined and limited duration. Licenses govern access to different types of data: (i) the current issues of journals; (ii) the electronic back issues of journals, i.e. the online back-files of born-digital journals; and (iii) the digital archives, i.e. the digitised versions of the back issues of printed journals.

In some cases the online backfiles are made freely available after a number of years.

As libraries do not own electronic journals, they cannot ensure their long-term accessibility or their preservation for the future.

This section addresses the question of how permanent access to and the long-term preservation of electronic scholarly journals, and thereby of research output, are currently being guaranteed.

8.1. ONLINE ACCESS TO JOURNAL BACK ISSUES

Publishers usually distinguish between back-files, i.e. the back issues of born-digital journals, and digital archives, i.e. the digitised version of print journal back issues.

They are considered different products that are sold separately, along different models and under different provisions. They are discussed in the two following sub-sections.

8.1.1. BORN-DIGITAL JOURNALS

Born-digital journal back-files are usually available as of 1995-1997, when publishers started producing electronic journals. Publishers have a variety of models for licensing access to their online back-files –some include access to some or all years in their current licence fee, some charge a separate fee, and some provide free access to all after a certain period (Cox and Cox, 2003). The provisions ensuring access to the licensed material after termination of the contract also vary across publishers. Alternatively, as e-print archives continue to grow, the question is raised whether they can provide permanent access to journal articles.

A variety of models for online access

A recent survey (Cox and Cox, 2003) has shown that out of the 149 respondents, 3% of the publishers made their backfiles freely available after a year or less and a further 6% after a longer period; the 66 small not-for-profit publishers in the survey stood out in this regard with 8% making backfile access free immediately, 6% after 6 months, 17% after a year, 6% in the next subscription year, and an additional 6% after a longer period. For example, 197 out of the 852 journals hosted by HighWire Press offer free back issues (and 30 offer entirely free contents)¹²⁹. If the delay is carefully determined (it may differ between disciplines), providing free access to backfiles appears not to undermine sales.

A variety of provisions for ongoing access to previously licensed material after termination of the contract

Licences usually distinguish between access to and preservation of the data after termination of the contract. There are many variations across contracts. After termination of the licence or if electronic subscription to a title is cancelled, publishers will generally provide:

- ongoing online access to the licensed materials, either from the publisher's site or from a third party, generally against payment of an annual fee¹³⁰, calculated along differing lines (e.g. Elsevier¹³¹: maintenance fee (US\$ 5,500 in 2005) + variable amount calculated on the number of full text article downloads

in the preceding 12 months (US\$ 0.25 per download in 2005); Springer¹³²: fee negotiated per title per year; Proquest : a one-time fee; Wiley¹³⁴, Blackwell Publishing¹³⁵, Project Euclid¹³⁶ : unspecified);

- and/or a copy of the subscribed materials, generally on CD-ROM, that can be mounted locally, and must sometimes be paid for. This is strictly a preservation provision, as the files are usually image files, not searchable as such, that would require from the recipients to set up the infrastructure with effective search and retrieval capabilities. Moreover, the files may not contain all links and other features and functionality associated with the online version available via the publisher's server. And eventually, as electronic content continues to grow, the quantity of CD-ROMs will become prohibitive and make this provision unsustainable¹³⁷.

The situation is more complex when access to journals is provided through aggregators, some of them creating electronic archives for permanent access (e.g. OCLC Electronic Collections Online¹³⁸), others referring to each journal publisher's policy (e.g. Ovid¹³⁹, HighWirePress¹⁴⁰).

Given the heterogeneity of the provisions, libraries cannot manage, nor guarantee, permanent access to formerly-subscribed electronic journals, as they did for print journals.

In an attempt to harmonise the situation, some associations or joint working parties regrouping representatives of publishers, universities and/or subscription agents, have established "model licences" for electronic journals¹⁴¹. Such model licences help to promote best practice regarding permanent access and preservation terms in contracts. They usually specify access and preservation conditions after termination of the contract, recommending that the publisher provides continued access to the licensee, either without charge or for a fee. Publishers argue that providing access to a specific range of journal issues to someone who is no longer an active subscriber requires extra sophistication and maintenance of the publisher's access control system, which justifies charging an annual fee.

A practical downstream solution for ongoing access has been set up with the LOCKSS program (Lots of Copies Keep Stuff Safe) which aims to provide a persistent access preservation system. Launched in April 2004, it enables libraries to store their e-journal content, using a caching system that keeps it available even after the subscription has been cancelled. Today some 80 libraries are using this open source program, and more than 60 publishers, including very large ones, have agreed to contribute to the LOCKSS program (although only 77 titles have been released to the system so far, content is expected to grow at a steady pace¹⁴²).

Potential role of e-print archives for providing permanent access to journal articles

E-print archives aim to provide open access to research articles, but they represent a fragmented and incomplete reproduction of journal issues: deposit is made on a voluntary basis, there is no systematic archiving, thus no guarantee of exhaustiveness.

Although they definitely improve access to research output, they do not currently provide a strategic global and comprehensive solution for permanent access to all *journals* content archives. Note that PubMed Central is an exception in this regard, as it aims to be a digital archive of *life science journal* literature and offers publishers to deposit all their journal content, allowing for a delay in the full text release after publication.

Thus, as self-archiving policies expand and their content grows, networked e-print archives might in the longer term offer an efficient alternative access to *journal* archives, as today's archived articles become tomorrow's stock. However it will require technology developments to allow for efficient search, retrieval, linking, citation analysis etc, across scattered archives.

Conclusions

At present, online access to previously licensed material mostly relies on the publishers' provisions to assure persistent access to content, from its own platform or from a third party. It also relies on the libraries' ability to pay the potential annual ongoing access fee. And in case publishers go out of business, access to journals ultimately relies on their long-term preservation strategy (see below).

As their content grows, open access e-print archives will improve the accessibility of past journal articles, though they do not provide a global and comprehensive solution ensuring permanent access to all journal archives.

8.1.2. DIGITIZED JOURNALS

Most major publishers have started large scale and costly programmes to digitise their collections of print journals, back to the very first issues¹⁴³. Publishers then make the digital archives available on their e-journal platform and generally license access separately at an additional cost. Of course, the availability of the digital archive is in itself a considerable benefit to users.

Some publishers hand over digitization of their journals back issues to third parties, as well as their subsequent online access provision.

JSTOR¹⁴⁴ is such an aggregator that aims to preserve and provide long-term access to digitized scholarly journals. Initially funded by the Mellon Foundation as a project, JSTOR has since been established as an independent not-for-profit organization in 1995 that is now self-sustaining. More than 320 publishers have agreed to have their journals available in JSTOR, representing 664 journals grouped in different collections. The "moving wall", i.e. the time period between the current journal issue and the content available in JSTOR, is specified by publishers in their license agreements with JSTOR, and generally ranges from three to five years. Through the moving wall, JSTOR seeks to avoid jeopardizing publishers' subscriptions and revenue opportunities from current and recent material. In JSTOR, access is licensed to collections of titles, multidisciplinary or discipline-specific; fees are established

according to the size of the institution and consist of an "Archive capital fee", paid one time intended to help underwrite the costs of digitizing new collections and to ensure that JSTOR can continue to migrate data and software into the future, and an annual access fee to support the provision of ongoing access to the archive and cover the addition of new volumes as each year is added to the archive.

Other retro-digitisation projects are being set up, partly or fully funded by public authorities.

PubMed Central (PMC), for instance, besides operating as an e-print archive, also acts as a journal archives aggregator. The US National Library of Medicine (NLM) offers publishers to scan the back issues of their journals deposited in PMC that are not already available in electronic form. The complete contents of the scanned issues are made available free in PMC. The full cost of scanning the back issues is covered by NLM, though some specific journals are sponsored by the Wellcome Trust and the U.K. Joint Information Systems Committee (JISC)¹⁴⁵.

In some countries, publicly-funded initiatives have emerged for digitizing scholarly journals and providing access to electronic archives in Social Science and the Humanities (SSH). As the high costs and the relatively smaller user community -when compared to STM journals- appear unattractive for private investments, some national public authorities have started to finance projects to digitize journal collections in SSH¹⁴⁶. Such projects are often led and realised by university libraries. In France, for instance, the SSH portal Persée¹⁴⁷, funded by the Ministry of State Education, Higher Education and Research, aims to digitise and provide free online access to back collections of a large range of selected journals in SSH. More specific projects are also being supported, such as NUMDAM, funded by CNRS to digitize French mathematical journals. Wider journal dissemination and availability is expected to promote French research output, increase its use and ensure its preservation. Although national public authorities in France are showing rising interest in providing online access to journals in SSH, other countries, such as Spain and Italy, are less reactive¹⁴⁸.

Platforms that provide electronic publishing services and online access to journals sometimes include retro-digitisation opportunities. Érudit, the Canadian publicly-funded portal¹⁴⁹, produces and provides online access to more than 50 journals in SSH. Current issues are accessible on subscription while back issues are *freely* available on basis of the “moving wall” principle. Érudit is intended to serve as an innovative means of promoting and disseminating the results of university research, and to contribute to the development of leading-edge expertise in academic-publication digital publishing. CAIRN¹⁵⁰, for instance, is a recently created platform providing electronic publishing and dissemination services for journals in social sciences and humanities in French, also offering retro-digitisation services. CAIRN has benefited from Érudit expertise and applies its business model while relying on a public-private partnership involving small publishers and institutional investors from the scientific and university community.

National libraries are also involved in large cultural heritage digitisation projects which often include ancient collections of journals (e.g. the Bibliothèque Nationale de France’s Gallica collection). The recent EC “i2010: Digital Libraries” Communication¹⁵¹ underlines the need for strategic discussion with stakeholders, reinforced coordination and further co-funding to digitise and preserve the European cultural heritage, and announces a forthcoming specific Communication about scientific information.

Many projects, European and other, are already providing useful recommendations and guidelines for digitisation standards and best practice, which improve interoperability between the initiatives.

Conclusions

Business models to digitise print journal collections and to provide online access to digital archives may vary according to the size of the publisher (number of journals and revenues), to the domain, to national policies, to the organisation and structure of the scholarly publishing market in various countries, and to language.

The following trends have been observed:

- Publishers with large portfolios tend to digitise their complete collections and to license access to digital archives directly from their e-journals platform;
- Small publishers, learned societies are more likely to subcontract digitisation of their back issues and the subsequent online access, agreeing on a moving wall, with non profit organisations like JSTOR (licensing access) and PubMed Central (providing open access);
- Publicly-funded initiatives and public-private partnerships are emerging to digitise and provide online access (free or charged) to quality SSH journals in non-English speaking countries, in order to promote their national research output.

Digitisation and online accessibility of scientific publications will face the same challenges as those underlined in the i2010: Digital Libraries Communication for digitising the cultural heritage: financial – digitisation is labour-intensive and costly, choices will have to be made (see for instance in Persée); organisational – sustained coordination effort at national and European level and new partnerships according to the existing publishing market organisation in different countries, i.e. public vs. private actors; legal – different implementations of the Directive 2001/29/EC on copyright; and technical – improve digitisation techniques and optical character recognition for languages other than English.

8.2. LONG-TERM PRESERVATION OF ELECTRONIC JOURNALS

Besides digitising and providing online access to digital archives in order to meet current research needs, publishers and information providers also face questions of the long-term preservation of electronic journals for the future generations.

Digital preservation is “a process by which digital data is preserved in digital form in order to ensure the usability, durability and intellectual integrity of the information contained therein”¹⁵².

8.2.1. TECHNICAL DEVELOPMENTS

Technically, digital preservation implies setting up strategies, such as migration and emulation. Thorough research and studies that have been and are still being conducted about digital preservation emphasize the complexity of digital objects and the many characteristics to be accounted for in a long-term preservation scheme. As underlined in the EC i2010:Digital Libraries Communication, at present there is little experience with digital preservation. Questions are raised that deserve more research and practical implementation regarding: the exact contents of a journal that need to be preserved (selection criteria); the long-term sustainability of digital preservation supports; standards for preservation processes and techniques –the XML mark-up language and the OAIS model (*Open Archival Information System*)¹⁵³ are promising; metadata about preservation –such as the scheme recently defined by OCLC and RLG¹⁵⁴; methods protecting the document authenticity; technical strategies for sustainable archiving systems.

Research should also be supported in closely related domains that have a direct impact on preservation, such as: the creation of user-friendly XML editing tools enabling authors to directly and easily create articles in XML format –avoiding the subsequent necessary XML conversion for optimal document preservation; the automatic generation of metadata for large amounts of documents to be preserved; the creation of interoperable quality archives to deposit documents in different places and improve preservation conditions.

8.2.2. ROLES AND RESPONSIBILITIES

In the print era, the long-term preservation of scholarly journals was ensured formally by the national library through legal deposit, as well as informally by many other institution and research libraries that subscribed to the journals. Today, the responsibilities for digital preservation of scholarly journals are not yet clearly

established. This is however a crucial issue for the future long-term availability of research results.

Very few publishers make public their long-term preservation policy. The most recent surveys of publishers (dated 2003)¹⁵⁵ show varying degrees of concern and involvement; while an EC study reports that most publishers deposit their digital publications offline (Oskamp, 2003), an ALPSP survey indicates that only 50% of the publishers have taken formal provisions for the long term preservation of their journals (Cox and Cox, 2003). The current provisions involve multiple and distributed responsibilities: own arrangements, agreement with national library or another library, deposit with other third parties (JSTOR, NLM-PubMed Central, LOCKSS, OCLC, INASP, Ingenta...).

Publishers' own arrangements

Digital preservation is costly and publishers expect return on investments. There is a risk that publishers will stop to ensure digital preservation when licensing access to digital archives will no longer be profitable.

It appears most likely that publishers will assure initial archiving and provide access to digital archives as long as publications bear a commercial value, while making formal arrangements for long term preservation with a third party.

National libraries and legal deposit

“Legal deposit is a statutory obligation that requires any organization, commercial or public, and any individual producing any type of documentation in multiple copies, to deposit one or more copies with a recognized national institution” (CEDARS, 2002). Such institutions aim to preserve the national cultural and intellectual heritage, and to provide access to it, although access is often limited from inside the premises of the institution. National legislations are progressively being adapted in different countries to include deposit of digital materials, for instance in Austria, Germany, Great Britain, Canada, Norway, Denmark, Italy, France (project)¹⁵⁶.

Besides these legal provisions, some national libraries have negotiated directly with publishers to set up procedures for voluntary deposits of their electronic journals: such initiatives are witnessed for example in the Netherlands (*Koninklijke Bibliotheek's* agreements with Elsevier, BioMedCentral, Kluwer, Taylor&Francis, Blackwell, Oxford University Press et Springer)¹⁵⁷, and in Germany (*Deutsche Bibliothek's* agreement with Springer-Verlag¹⁵⁸). The *Conference of European National Librarians and the Federation of European Publishers* have published a code of practice intended to facilitate the voluntary deposit of electronic publications¹⁵⁹.

National and university libraries and other associations have very early joined forces to develop preservation strategies and reduce costs, and continue to collaborate in specific projects¹⁶⁰. These initiatives are to be placed in the larger context of Web archiving, for which many experiments are taking place notably in Australia, Germany, the Netherlands, France, Austria, Sweden and Denmark¹⁶¹.

However many problems remain for the practical implementation of electronic resources preservation such as determining the criteria to select what deserves to be preserved, and bringing the necessary investments in material and human resources.

It should be noted that the institutions offering publishers to deposit electronic publications, be it legal or voluntary, aim at ensuring the long-term preservation of the resources and do not provide *remote* online access to the digital archives.

Other third parties

As mentioned earlier in this report, JSTOR is a non-profit organisation ensuring the preservation of digitized scholarly journals and licensing access to the collections. In its business plan, it has provided for an Archive Capital Fee to cover data and software migration into the future.

E-Archive¹⁶² is a promising recent initiative, originally launched by JSTOR and now incubated within Ithaka¹⁶³, which aims to provide the infrastructure to preserve scholarly literature produced in electronic form and to ensure that these materials remain accessible to future scholars, researchers, and students. A critical

aspect of this undertaking is devising a fair way to share the costs and overcome the 'public good' economic hurdle which has thus far precluded large scale e-archiving efforts. In pursuing this mission, E-Archive has adopted a system-wide perspective, taking into account the sometimes conflicting needs of libraries, publishers and scholars.

The specific objectives of E-Archive are:

- To ensure that the record of electronically-published scholarship is reliably maintained.
- To create a production-level archiving solution and to fairly distribute costs among academic institutions, libraries, publishers and government agencies.
- To enable libraries to realize cost savings through more aggressive migration to e-resources with the assurance of the future availability of important e-resources.

Development of the key infrastructure elements is still underway. The resulting E-Archive service and business model should balance the interests of publishers and libraries, though clearly focussing on preservation purposes, access being provided "under limited circumstances". It is therefore complementary to JSTOR that aims to license access to the journals contents.

PubMed Central, mentioned previously, also aims at ensuring the long term preservation of life science journals, and provides open access to the archived journals.

Conclusions

Publishers' current provisions for the digital preservation of their journals involve multiple and distributed responsibilities: own arrangements, agreements with national library or other libraries, and deposit with other third parties. As underlined in the EC "i2010: Digital Libraries" Communication, the challenges for digital preservation are similar to the ones for digitisation:

- *financial* –the real costs of long-term digital preservation are not clear, choices have to be made as to which material should be preserved;
- *organisational* –coordination, new business models and collaboration are needed to balance the interests of all stakeholders;

- *technical* –preserve high volumes of rapidly changing information;
- *legal* –legal deposit schemes provide for preservation but not for remote availability.

At present, the digital preservation archives that are being set up ensure the long-term preservation of electronic journals, as a service for publishers (legal and voluntary deposit, E-Archive), but do not in general license remote online access. JSTOR and PubMed Central are exceptions, providing additional services like journal digitisation and online access. Such non-profit organisations combining the different services might appear more cost-effective. The E-Archive business model is likely to be promising as well in balancing the interests of the different parties. It should be underlined, however, that public institutions are the only organisations that are able to guarantee the long-term preservation of scientific publications (just as of cultural heritage) and its accessibility to future generation.

8.3. CONCLUSIONS

Publishers have a variety of business models for providing access to their online back-files –some include access to some or all years in their current licence fee, some charge a separate fee, and some provide free access to all after a certain period. Today, online access to previously licensed material mostly relies on the publishers’ provisions to assure persistent access to content, from their own platforms or from a third party. It also relies on the libraries’ ability to pay the potential annual ongoing access fee. As self-archiving policies expand and the contents of open access e-print archives grow, it will improve the accessibility of past journal articles, though e-print archives do not provide a global and comprehensive solution ensuring permanent access to all journal archives.

In order to guarantee perennial access to scholarly journals digital archives, given the heterogeneity of the publishers’ current provisions, one could promote the creation of not-for-profit long-term preservation journal archives, under the aegis of some public institutions, which balance interests among publishers, libraries, and scholars.

More particularly:

- Promote economic models for journal digital archives that are based on the notion that participating institutions can save resources over the long term by contributing to the costs of a centrally managed archive. E.g. the JSTOR participation fee, comprising a fixed archive capital fee and an annual access fee.
- Encourage the “moving wall” model, i.e. allowing for a time period between the current journal issue and the content available in the archive. This prevents jeopardizing publishers’ subscriptions and revenue opportunities from current and recent material, while also enabling libraries and researchers to rely on archives, providing both preservation and access to journals after a reasonable period of time.
- Encourage public funding and public-private partnerships where there is little commercial investment in the creation of journals digital archives, especially for quality European journals in SSH.
- Where legal deposit for digital material exists, develop business models for legal-deposit libraries to provide remote online access to their journals digital archives, on a non-profit basis (like JSTOR, self-sustaining organisation) or for free (like PMC, publicly-funded organisation); such business models would provide these publicly-funded libraries (such as legal deposit libraries) with return on investments (financially or through wider public access to scientific heritage) and rationalize the overall process of digital archives creation and access provision, making the preservation efforts cost-effective.
- Investigate the feasibility/relevancy to create a European non-profit journal preservation organisation (“JSTOR-like”) and to set up other subject-based archives in relevant domains like life sciences (e.g. a “European PMC”), under the aegis of public institutions responsible for perennial access. Such journal digital archive organisations would be aimed at quality European journals and provide (i) publishers with long-term preservation services, where national legal deposit for digital material is not yet effective, and (ii) researchers and libraries with permanent access to these journal archives, under a non-profit business model to be determined. Journal archives are intended for preservation and access to the

“stock” of knowledge, whilst open access e-print archives are intended for improving dissemination and access to recent research outputs and meeting the research community needs and practices, such as fast dissemination pre-prints; central access to the scientific production of a research institution, be it articles, thesis, working paper, conferences, etc.; providing potential tools for research monitoring and policy making at institutional, national and European level. In some disciplines, like biomedical and life sciences, where only peer-reviewed articles should be disseminated, it might be adequate to create a single archive for preserving journals and for self-archiving articles (e.g. a “European PMC”).

- Determine the standards under which the journal archives must be accessible and ensure their compatibility with the open self-archiving repositories collecting the authors’ articles in order to guarantee interoperability across all scientific publication archives.
- Set up a portal as a central access point to digital journal and article archives, whether they are stored in subject-based dedicated digital archives (European and US PMC, JSTOR etc) or in legal-deposit libraries, or in open access e-print repositories.

It is expected that the forthcoming EC “i2010: Digital Libraries” Communication focusing on scientific information will address these issues of journal digitization, online availability and preservation. The type of challenges is similar to those for cultural heritage, though the market and the stakeholders differ in several respects. The scientific information sector will directly benefit from research in digitisation and digital preservation funded under the Information Society Technologies programme and the Seventh Framework Programme. The network of Centres of Competence, intended to become the cornerstone of European digitisation and preservation, could also contribute in transferring knowledge and expertise to scholarly publishing actors.

¹³⁵ URL: <http://www.blackwellpublishing.com/license/license.pdf>

¹³⁶ URL: http://projecteuclid.org/documents/ProjectEuclid_SubscriberLicense.pdf

¹³⁷ Watson (2005)

¹³⁸ OCLC, Ensuring ongoing access. URL: <http://www.oclc.org/middleeast/en/electroniccollections/about/ongoing/> and OCLC, Archiving solution. URL: <http://www.oclc.org/middleeast/en/electroniccollections/archiving/>

¹³⁹ URL: http://www.ovid.com/site/products/journals_archive_policy.jsp?top=2&mid=3&bottom=7&subsection=12

¹⁴⁰ URL: <http://highwire.stanford.edu/institutions/archiving.dtl>

¹⁴¹ See for instance http://www.alpssp.org/http_licens.htm

¹⁴² See <http://www.lockss.org/about/titles.htm>

¹⁴³ Elsevier for instance invested over 32 million Euros in scanning and digitizing the back issues of its 1,800 journals. (Reed Elsevier, 2005)

¹⁴⁴ JSTOR, the scholarly journal archive. URL: <http://www.jstor.org/>

¹⁴⁵ The NML Back issue digitization project <http://www.pubmedcentral.nih.gov/about/scanning.html>

¹⁴⁶ Minon and Chartron (2005)

¹⁴⁷ URL: <http://www.persee.fr/>. The consortium consists of the University Lumière Lyon 2, the Maison de l’Orient et de la Méditerranée - Jean Pouilloux (Institute of the East and the Mediterranean - Jean Pouilloux) and the University of Nice-Sophia Antipolis, in collaboration with the Ecole Normale Supérieure de Lyon (élite humanities university).

¹⁴⁸ See for instance Minon and Chartron (2005).

¹⁴⁹ URL: <http://www.erudit.org/>. Funded by Fonds québécois de recherche sur la société et la culture, Fonds de l’Autoroute de l’information, Ministère du Développement économique et régional, Fonds québécois de recherche sur la nature et les technologies and Fonds de recherche en santé du Québec.

¹⁵⁰ URL: <http://www.cairn.info>

¹⁵¹ EC Information Society. i2010: Digital Libraries, 2005.

¹⁵² Russell (1999)

¹⁵³ Research Libraries Group (2002).

¹⁵⁴ OCLC/RLG Working group on preservation metadata (2002).

¹⁵⁵ Muir (2004), Oskamp (2003), Jones (2003), Cox and Cox (2003)

¹⁵⁶ See Preserving Access to Digital Information PADI URL: <http://www.nla.gov.au/padi/topics/67.html>

¹⁵⁷ Elsevier Science <http://www.kb.nl/nieuws/2002/elsevier.html> (20 augustus 2002) ; Biomed Central <http://www.kb.nl/nieuws/2003/kb-biomed.html> (17 september 2003) ; Kluwer Academic Publishers (19 mei 2003) ; Taylor & Francis (1 november 2004); Blackwell (22 juni 2004) ; Oxford University Press (13 oktober 2004) ; Springer (6 april 2005).

¹⁵⁸ Submission of online publications to Die Deutsche Bibliothek, février 2005. URL: http://deposit.ddb.de/netzpub/web_abgabe_np_gesamt_e.htm

¹⁵⁹ CENL/FEP (1999)

¹⁶⁰ See for example Preserving Access to Digital Information PADI: <http://www.nla.gov.au/padi/>; NEDLIB <http://www.kb.nl/coop/nedlib>; PREMIS Working Group, 2004 ; CIDL Canadian Initiative on Digital Libraries <http://www.collectionscanada.ca/cidl/>; CEDARS <http://www.leeds.ac.uk/cedars/>; NESTOR http://www.langzeitarchivierung.de/modules.php?op=modload&name=PagEd&file=index&page_id=23; NDIIPP National Digital Information Infrastructure and Preservation Program funded by the US Library of Congress and the National Science Foundation http://www.digitalpreservation.gov/about/pr_050405.html; JISC program Supporting Digital Preservation and Asset Management in Institutions http://www.jisc.ac.uk/index.cfm?name=programme_404

¹⁶¹ Australia (PANDORA): <http://pandora.nla.gov.au/index.html>; Germany (DEPOSIT.DDB.BE): http://deposit.ddb.de/index_e.htm; the Netherlands (Edepot): <http://www.kb.nl/dnp/e-depot/dm/dm-en.html>; France: http://www.bnf.fr/pages/infopro/depotleg/dli_intro.htm; Austria (AOLA): <http://www.ifs.tuwien.ac.at/~aola/>; Sweden (Kulturarv3): <http://www.kb.se/kw3/ENG/Default.htm>; Denmark (Netarchive.dk): <http://www.netarchive.dk/index-en.htm>. See complete list on PADI web site Web archiving, URL: <http://www.nla.gov.au/padi/topics/92.html>.

¹⁶² <http://www.ithaka.org/e-archive/index.htm>

¹⁶³ Ithaka is a newly formed not-for-profit organization with a mission to accelerate the productive uses of information technologies for the benefit of higher education around the world. <http://www.ithaka.org/>

¹²⁹ Watson (2005)

¹³⁰ Ongoing access is sometimes provided free of charge as long as the customer subscribes to a least one electronic journal on the publisher’s platform

¹³¹ URL: http://www.info.sciencedirect.com/licensing_options/policies/index.shtml

¹³² URLs: <http://www.springeronline.com/sgw/cda/frontpage/0,11855,5-117-2-130176-0,00.html#anchor6> (question n°3) et <http://www.springeronline.com/sgw/cda/frontpage/0,11855,5-117-2-128726-0,00.html>.

¹³³ URL: <http://www.proquest.com/terms.shtml>

¹³⁴ URL: http://www3.interscience.wiley.com/spreadsheet_documents/sample_ballicense.pdf

9. STANDARDS AND INTEROPERABILITY

Standards are a key element in the publishing system; as different open and proprietary applications are being developed by different actors, the use of standards is crucial to enable and facilitate data exchange and communication on the network, and thus definitely improve dissemination and access to scholarly publications.

Their use at every stage of the publishing process –creation, description, dissemination, preservation- will ensure that the user can search, view and print the article whatever its hosting platform, that scientific publications are widely visible and searchable from a variety of service providers (be they free or charged), and that the scientific literature remain accessible for the future. We underline future developments that could be supported by research funding programmes:

Metadata provide structured information that is crucial for the digital object exchange, use and management. Though some descriptive metadata standards are well established in the libraries and publishers sectors, their specificity hinders their efficiency in enlarged scientific search and retrieval applications through the internet. On the other hand, the widely used and supported Dublin Core metadata set appears too general to provide fine-grained search results, adapted to scientific information search practices. This calls for a standard qualified version of the Dublin Core for scientific outputs to be defined. Standards for other types of metadata -notably for technical metadata, rights management metadata and preservation metadata- should also be further investigated and used. Metadata are essential elements to the future of scientific publishing and to the development of new forms of scholarly information dissemination such as e-print repositories. The efficiency of harvesting protocols, such as the OAI-PMH, and of search engines and services built on the harvested

metadata repositories relies on the quality and adequate granularity of the metadata.

The **XML format** appears as the main thread in the whole publishing process, i.e. creation, description, dissemination and preservation. Developing software applications for the creation of XML documents would highly simplify the publishing process and help its wider use at all stages of the process (e.g. a Wysiwyg (What you see is what you get) editing tool). Furthermore, in order to take full advantage of XML throughout the publishing process, additional research and developments would be welcome, for instance, to define a standard XML schema to structure the full text, to specify a qualified Dublin Core adapted to scholarly documents, to automatically create metadata from the full text. Further developments of the XML format will improve and accelerate the overall publishing process.

At present, searching publications by name of author or institutions appears difficult due to homonyms, spelling errors, the use of acronyms, reduced or translated institution denominations, etc. A **directory of European researchers and research institutions** enabling their unique identification (e.g. through unique identifier) would improve search efficiency as well as the monitoring of research outputs and publications in Europe.

Persistent digital object identifiers should be preferred to URL. In the publishing sector, the DOI (Digital Object Identifier) system, recently accepted for standardisation within ISO, was developed and is managed by the International DOI Foundation (IDF)¹⁶⁴. CrossRef is the largest DOI registration agency that covers scholarly and professional research content. As IDF and CrossRef associations are directed by the main commercial scientific publishers, control over DOI allocation has raised concerns as an internet governance issue in the WSIS Civil Society working group¹⁶⁵. This group might recommend the creation of an inclusive multi-stakeholder partnership under the aegis of the United Nations to implement a free and public digital object identifier system. The opportunity to create such an agency should be further investigated.

Besides advanced search functionalities, **linking technologies** are being integrated in article databases such as hyperlinks, which launch a new search when clicking on the item. Furthermore, various linking technologies enable data hosted on different servers to be related:

- Reference linking, provides links from the bibliographic references of an article to the full-text articles of other platforms. CrossRef provides publishers with such a cross-publisher citation linking system based on DOIs, allowing a researcher to click on a reference citation on one publisher's platform and link directly to the cited content on another publisher's platform, subject to the target publisher's access control practices.
- The OpenURL standard (*Open Uniform Resource Locator*¹⁶⁶) provides an extended and dynamic open linking technology, allowing a user to navigate across different platforms, be they reference, article or full-text resources, which are relevant to his search topic. An increasing amount of data providers, including CrossRef, are making their database OpenURL compliant to facilitate navigation to their content.
- Other linking systems provide information on citation counts and links to citing articles ("cited-by"). Such added-value functionalities are found in subscription-based services (e.g. Thomson ISI Web of Science, Elsevier Scopus) and more recently in Google Scholar which automatically analyzes and extracts citations and presents them as separate results, even if the documents they refer to are not online.

These technologies and services are improving cross-platform search and navigation to scholarly publications available on the internet. However improvements still have to be achieved, particularly at the description and indexing level, in order to reach the high-performance search level of dedicated abstract & indexing databases.

The exchange of data and their integration in large searchable repositories present today major ICT challenges.

Information retrieval protocols are being developed along two approaches:

- federated search, i.e. searching different systems simultaneously and grouping the results displayed to the user¹⁶⁷;
- harvesting metadata, i.e. gathering metadata from different systems into a central repository that can be searched. The **OAI-PMH**¹⁶⁸ (*Open Archives Initiative Protocol for Metadata Harvesting*) protocol is based on HTTP and XML, transforming metadata into the Dublin Core format for harvesting. Search applications can then be built on the resulting repository as well as science-dedicated versions of Web search engines (e.g. Google Scholar, Yahoo!, Scirus...). However the reliance on the Dublin Core that brings poor metadata and the disparity of the metadata harvested from heterogeneous repositories bring forward a challenge in terms of search and retrieval efficiency (vocabulary control, display...).

Interoperable standard protocols, such as the OAI-PMH, should be promoted to enable search across platforms and allow new user-oriented services to be built and provided to the research community.

As a lot of standards and new technologies have been developed, **collaboration and standardization** are essential to increase **interoperability** across systems and platforms in order to: facilitate the efficient dissemination of scholarly content, improve their visibility, open up access to a wide range of digital materials and ensure their long-term preservation (e.g. promote existing standards through metadata registries and create crosswalks between metadata standards)

The recently created *Knowledge Exchange Network* grouping four major organisations¹⁶⁹ promoting the use of ICT within research and higher education and already funding related projects might play a central role in this respect. The Network aims to focus on "the development of joint strategies and the identification of common or compatible best practices, the ultimate aim being to provide scientists and students with user-friendly systems that cross national boundaries and provide them with

access to scientifically relevant information”¹⁷⁰. Supporting collaboration, promoting good practices and the use of standards and interoperability protocols in the publishing process might fit into its mission.

At the European level, the scientific publishing sector will benefit from research funded under the EC Information Society Technologies programme as well as the eContentPlus programme that supports projects improving the accessibility and usability of European cultural and scientific content. For instance, achieving interoperability between national digital collections and services (e.g. through common standards) and facilitating access and use of the material in a multilingual context are core objectives that will serve the scholarly publishing sector as well. The forthcoming “i2010: Digital Libraries” Communication on scientific information could focus on the standard and interoperability issues specific to scholarly publishing and support the technical developments underlined in this section.

¹⁶⁴ URL: <http://www.doi.org>

¹⁶⁵ URL: <http://www.wsis-si.org/DOL/index.html>

¹⁶⁶ URL: <http://library.caltech.edu/openurl>

¹⁶⁷ E.g. Z39.50 and the more promising SRW (Search and Retrieve Web Service), relying on XML and HTTP (URL: <http://lcweb.loc.gov/z3950/agency/>; URL: <http://www.loc.gov/z3950/agency/zing/srw/>)

¹⁶⁸ URL: <http://www.openarchives.org> and <http://www.oaforum.org/index.php>

¹⁶⁹ It groups together the Joint Information Systems Committee (JISC) from the UK, the Dutch SURF Foundation, Denmark's Electronic Research Library (DEF) and the German Research Foundation (DFG).

¹⁷⁰ CORDIS RTD-News RCN :24168 « Countries work together to promote use of ICT within research », 19/07/2005 URL: http://dbs.cordis.lu/cgi-bin/srchidadb?CALLER=NHP_EN_NEW S&ACTION=D&SESSION=&RCN=EN_RCN_ID:24168

10. RECOMMENDATIONS

This section puts forward a set of policy recommendations for improvements in the scientific publication process. These recommendations follow the analysis contained in earlier sections.

Section 3 has documented the very significant price differences – from 1 to 3 between non-profit and for-profit publishers for comparable-quality journals (or articles). This dispersion is indicative of market power, as is the positive link between prices and number of citations. Since citations should be positively correlated with circulation, and thus with lower average costs (given the scale economies, due to high ‘first-copy costs’), this positive link is a sign of ‘value-based pricing’ rather than ‘cost-based pricing’, and thus again of market power.

While it is important to stress the societal value of the existing publication system, it is also important to acknowledge the societal cost linked to high journal prices, in financial terms for public budgets, but also in terms of limits on the dissemination of knowledge and therefore of further scientific progress. This can be particularly unfortunate at a time where technological possibilities due to the Internet have the potential to dramatically raise dissemination possibilities.

Consequently, policies should make sure that the market is sufficiently competitive and ‘dissemination-friendly’. In particular, they should address the need to:

- (i) enhance access to research output;
- (ii) prevent strategic barriers to entry and to experimentation.

The first five recommendations concentrate on access issues, which are currently the subject of significant policy discussions.

RECOMMENDATION A1. GUARANTEE PUBLIC ACCESS TO PUBLICLY-FUNDED RESEARCH RESULTS SHORTLY AFTER PUBLICATION

Research funding agencies have a central role in determining researchers’ publishing practices. Following the lead of the NIH and other institutions, they should promote and support the archiving of publications in open repositories, after a (possibly domain-specific) time period to be discussed with publishers. This archiving could become a condition for funding.

The following actions could be taken at the European level: (i) Establish a European policy mandating published articles arising from EC-funded research to be available after a given time period in open access archives, and (ii) Explore with Member States and with European research and academic associations whether and how such policies and open repositories could be implemented.

RECOMMENDATION A2. AIM AT A ‘LEVEL-PLAYING FIELD’ IN TERMS OF BUSINESS MODELS IN PUBLISHING

There is a central role for education and research funding authorities in the shaping of new models for publishing and communicating research results. They should be aware that the rules governing education and research budgets have strong implications for the viability of various business models. At this point, it seems desirable to allow for experimentation and competition between various possible business models, which means allocating money to libraries to subscribe to reader or library-pay journals but also to authors to pay for publication costs in author-pay journals, and to researchers in the reader-pay model. Establishing relative priorities in this respect should become a key policy debate.

RECOMMENDATION A3. ‘EXTENDED QUALITY’ RANKINGS OF SCIENTIFIC JOURNALS

This recommendation aims at raising researcher awareness of journal quality beyond scientific quality, *stricto sensu*. While scientific quality, approximated for example by citation counts, should remain the dominant criterion, dimensions

related to the **quality of dissemination** (self-archiving authorisation, publisher archiving provisions, copyright provisions, abstracting and indexing services, reference linking, etc.) could be tracked explicitly and possibly valued by research funding bodies. There could be an impetus from public authorities at the European level for such an initiative, which would naturally induce publishers to stress good practices in these dimensions.

RECOMMENDATION A4. GUARANTEE PERENNIAL ACCESS TO SCHOLARLY JOURNAL DIGITAL ARCHIVES

Given the heterogeneity of the publishers' current provisions, promote the creation of not-for-profit long-term preservation archives, which balance interests among publishers, libraries, and scholars. More particularly (i) Promote business models for legal-deposit libraries to allow remote online access to their journals digital archives, therefore providing them with return on investments and making the preservation efforts cost-effective; (ii) Investigate the feasibility/desirability of the creation of a European non-profit journals preservation organisation ("JSTOR-like") and of other subject-based archives in relevant domains; (iii) Determine the standards under which archives must be accessible and set up a portal as a central access point to digital journals and articles.

RECOMMENDATION A5. FOSTER INTEROPERABLE TOOLS TO IMPROVE KNOWLEDGE VISIBILITY, ACCESSIBILITY AND DISSEMINATION

This could be achieved by (i) supporting research and development on interoperability issues, notably on metadata to improve scientific information search and retrieval efficiency and on the XML format to improve and accelerate the overall publishing process, and by (ii) promoting the wide implementation of linking technologies, especially the open standard OpenURL, and of interoperable standard protocols, especially the OAI-PMH that enables metadata harvesting and searching across different platforms. Both developments could

be taken into account by the European Commission in its e-infrastructure building strategy for the European Research Area (involving DG-Information Society R&D funding programmes and the forthcoming "i2010: Digital Libraries" Communication on scientific information).

The next three recommendations concern the need to prevent strategic barriers to entry and to experimentation and also excessive concentration.

RECOMMENDATION B1. PROMOTE PRO-COMPETITIVE PRICING STRATEGIES

The key issue identified in terms of market access concerns pricing policies, and more particularly the lock-in effect associated to 'Big deals'. Specifically, the limited savings libraries obtain for net subscription cancellations does make it hard for newcomers to have access to library budgets. The following simple rules (that could be promoted by the European authorities) would avoid some of the long term negative effects of big deal contracts on entry and competition: (a) The price of the electronic access should not depend on the *historical* number of print subscriptions; (b) Prices should be related to transparent indicators, like usage or the number of faculty, students, etc., as is the case with JSTOR for instance; (c) Libraries should have the possibility to choose among variable dimension bundles, and compose their preferred bundle. Therefore, journals in a bundle should also be priced individually, and prices of bundles should ideally be made public; (d) Finally, note that overall usage has been on the rise thanks to the Internet, and can be expected to keep growing at least for some time. One should avoid having prices increase with such usage as long as publishing costs do not increase as a result of this rise in usage.

RECOMMENDATION B2. SCRUTINIZE FUTURE SIGNIFICANT MERGERS

The market has become more concentrated due in part to acquisitions by large for-profit publishers, and some of the price increases

can be traced back to these mergers, though the largest firm controls less than 30% of the overall market (market shares are however higher in some scientific fields). It has been shown that publishers with large journal portfolios have an incentive to set higher prices. This indicates that further acquisitions by large publishers should be scrutinized by the relevant European authorities.

RECOMMENDATION B3. PROMOTE THE DEVELOPMENT OF ELECTRONIC PUBLICATIONS

Eliminate the unfavourable tax treatment of electronic publications, by (i) either applying a reduced VAT rate to all types of scientific information, whether print or electronic; (ii) or, given the political difficulty of implementing this solution which requires unanimity of Member States, by introducing a tax refund mechanism for research institutions, as is already the case in Sweden and Denmark. Though the European Commission could play a role in supporting and promoting either solution, this decision is ultimately left to the Member States.

Encourage public funding and public-private partnerships where there is little commercial investment in the creation of journals digital archives, especially for quality European journals in Social Science and Humanities. Such initiatives require further investigation of the structure and organisation of publishing markets for SSH journals which are quite different between countries, especially the distribution among private and public actors.

Finally, the last two recommendations stress the need for further discussions and study concerning this important market.

RECOMMENDATION C1. SETTING-UP AN ADVISORY COMMITTEE

Discussions with all the stakeholders during the study made it clear that regular contacts are necessary, since the industry's practices are moving very fast, and will keep doing so in the future. We advise to set up a committee composed of the various interested parties:

publishers, librarians, funding bodies, authors and researchers, who should be responsible for observing practices, meeting (say once or twice a year) to discuss and recommend changes if need be, and reporting the results of the discussions to the Research (and possibly the Competition) DG's of the European Commission.

RECOMMENDATION C2. FURTHER INVESTIGATION

This study is obviously not exhaustive. Here are some topics where further investigation could be commissioned:

- A first important topic concerns the evolution of **copyright provisions**, which we address only briefly in this report. While publishers have become more permissive over time, in particular in terms of the posting of published material on individual web pages, it would be good to investigate precise legal solutions that would provide legal certainty to authors, but also potentially to other parties, in terms of dissemination of published material.
- A second topic concerns the **economic analysis of alternative forms of dissemination**: for example, the feasibility/desirability of alternative publishing business models (pay-per-download, author-pay systems, hybrid systems) and of the unbundling of certification and dissemination; and the long-term sustainability of open repositories.
- Finally, a third topic concerns **technological developments**: Research could be supported for example on interoperability issues and on the specifics of long-term preservation issues.

11. REFERENCES

11.1. ECONOMIC THEORY

- Adams, W. J. and J. L. Yellen (1976), Commodity bundling and the burden of monopoly, *Quarterly Journal of Economics* 90(3), 475-498.
- Aghion, P. and P. Howitt (1998), *Endogenous Growth Theory*, Cambridge, Mass.: MIT Press.
- Armstrong, M. (1996), Multiproduct Nonlinear Pricing, *Econometrica*, 64(1), 51-75.
- Bakos, Y. and E. Brynjolfsson (1999a), Bundling information goods: pricing, profits, efficiency, *Management Science*, 45(12), 1613-1630.
- Bakos, Y. and E. Brynjolfsson (1999b), Bundling and competition on the internet, *Marketing Science*, 19(1), 63-82.
- Baye, M. R. and J. Morgan (2001), Information Gatekeepers on the Internet and the Competitiveness of Homogeneous Product Markets, *American Economic Review*, 91(4), 454-474.
- Brown, K. and P. Alexander (2004), Bundling in the cable television: A pedagogical note with a policy option, *International Journal of Media Management* 6 (384), 162-167.
- Caillaud, B. and B. Jullien (2003), Chicken and Egg: Competition among Intermediation Service Providers, *RAND Journal of Economics*, 34(2), 521-552.
- Dewatripont, M. and P. Legros (2000), Mergers in Emerging Markets with Network Externalities: The Case of Telecoms, in *Le nouveau modèle économique européen* (vol. 2), eds. P. Magnette and E. Remacle, Brussels, Belgium: Editions de l'Université Libre de Bruxelles.
- Dobson Consulting (1999), Buyer power and its impact on competition in the food retail distribution sector of the European Union, prepared for the European Commission, DGIV Study contract No. IV/98/ETD/078.
- Ellison, G. and D. Fudenberg (2003), Knife-Edge or Plateau: When Do Market Models Tip?, *Quarterly Journal of Economics*, 118(4), 1249-1278.
- Fay, S. and J. K. Mackie-Mason (1999), Competition between firms that bundle information goods, Working Paper, University of Michigan.
- Ginsburgh, V. and I. Zang (2005a), Bundling by Competitors and the Sharing of Profits, *CORE Discussion Paper 2005/7*.
- Ginsburgh, Victor and Israel Zang (2005b), Notes on Bundling when Customers Valuations of Products are Dependent, mimeo Ecares.
- Laffont, J.-J. and J. Tirole (1993), *A theory of Incentives in Procurements and Regulation*, Cambridge, Mass.: MIT Press.
- Mc Afee, R. P., J. McMillan, M. D. Whinston (1989), Multiproduct Monopoly, Commodity Bundling, and Correlation of Values, *Quarterly Journal of Economics*, 114, 371-384.
- Nalebuff, B. (2004), Bundling as an entry barrier, *Quarterly Journal of Economics*, 119, 159-188.
- Rochet, J.-C. and J. Tirole (2004), Two-Sided Markets; An Overview, mimeo IDEI, Toulouse.
- Salinger, M. A. (1995), A Graphical Analysis of Bundling, *Journal of Business*, 68, 85-98.
- Schmalensee, R. L. (1984), Gaussian Demand and Commodity Bundling, *Journal of Business* 57, S211-S230.
- Schumpeter, J. A. (1934), *The Theory of Economic Development*, Cambridge, Mass.: Harvard University Press.
- Shapiro, C. and H. R. Varian (1998), *Information Rules: A Strategic Guide to the Network Economy*, Cambridge, Mass.: Harvard Business School Press.
- Shy, O. (2001), *The Economics of Network Industries*, Cambridge, UK: Cambridge University Press.

- Smith, A. (1776), *An Inquiry into the Nature and Causes of the Wealth of Nations*, London: Methuen and Co.
- Stephan, P. E. (1996), The Economics of Science, *Journal of Economic Literature*, 34(3), 1199-1235.
- Chressanthis, G. A. and J. D. Chressanthis (1994), The Determinants of Library Subscription Prices of the Top-Ranked Economic Journals: An Econometric Analysis, *The Journal of Economic Education*, 25(4), 367-382.

11.2. SCIENTIFIC PUBLISHING

- Baldwin, C. (2004), *What do Societies do with their Surpluses?*, West Sussex, UK: The Association of Learned and Professional Society Publishers and Blackwell Survey.
- Ball, D. and J. Pye (2000), Library Purchasing Consortia: The UK Periodicals Supply Market, *Learned Publishing*, 13(1), 25-35.
- Bergstrom, T. C. (2001), Free Labor for Costly Journals?, *Journal of Economic Perspectives*, 15(4), 183-198.
- Bergstrom, C. T. and T. C. Bergstrom (2004), The Costs and Benefits of Library Site Licenses to Academic Journals, *Proceedings of the National Academy of Sciences of the United States of America*, 101(3), 897-902.
- Bevan, S., P. Dalton and A. Conyers (2005), How Usage Statistics Can Inform National Negotiations and Strategies, *Serials*, 18(2), 116-123.
- Binman, J., R. Kirby and K. Apr (1997), *Comparative Prices of Maths Journals*, mimeo, Berkeley.
- CEDARS (2002), *Cedars guide to intellectual property rights*.
URL: <http://www.leeds.ac.uk/cedars/guideto/ipr/>
- CENL/FEP (1999), *Statement on the development and establishment of codes of practice for the voluntary deposit of electronic publications*, Conference of European National Librarians/Federation of European Publishers.
URL: <http://www.bl.uk/gabriel/fep/>
- Competition Commission (2001), *Reed Elsevier Plc and Harcourt General, Inc: A report on the proposed merger*, Report for the Secretary of State for Trade and Industry, UK.
- Cornell (1998), *Journal Price Study of Core Agricultural and Biological Journals*, Cornell University, Ithaca, New-York.
- Cox and Cox (2003), *Scholarly Publishing Practice : the ALPSP report on academic journal publishers' policies and practices in online publishing*, West Sussex, UK: ALPSP.
- Crow, R. (2002), *The Case for Institutional Repositories: A SPARC Position Paper*
URL: http://www.arl.org/sparc/IR/IR_Final_Release_102.pdf (SPARC, 2002).
- Crow, R. and H. Goldstein H. (2003), *Guide to Business Planning for Converting a Subscription-based Journal to Open Access*, Open Society Institute.
- CSFB (2004), *STM Publishing, Evolving threats to STM Publishing*, Credit Suisse First Boston.
- CSFB (2004), *Reed Elsevier. Scopus - Moving the Goalposts* , Credit Suisse First Boston.
- CUL (2004), Davis, P., T. Ehling, O. Habicht, S. How, J. M. Saylor and K. Walker, *CUL Open Access Publishing Task Force Report*, Presented to the Cornell University Library Management Team
URL: <http://techreports.library.cornell.edu:-8081/Dienst/UI/1.0/Display/cul.lib/2004-3>
- Cummings, A. M., M. L. Witte, W. G. Bowen, L. O. Lazarus, and R. H. Ekman (1992), *University Libraries and Scholarly Communication*, A Study Prepared for The Andrew W. Mellon Foundation, Washington, DC: Association of Research Libraries.

- De Heer, D. (2005), How Elsevier Uses Usage Information: The Inside Scoop, *Library Connect*, Elsevier.
- Deutsche Forschungsgemeinschaft (2005), Publikationsstrategien im Wandel? Ergebnisse einer Umfrage zum *Publikations- und Rezeptionsverhalten unter besonderer Berücksichtigung von Open Access*
URL:http://www.dfg.de/dfg_im_profil/zahlen_und_fakten/statistisches_berichtswesen/open_access/download/oa_ber_dt.pdf (2005)
- DG COMP (2003), *Report on merger CINVEN/CANDOVER / BERTELSMANN*.
- Dirkmaat, J. (2002), Comments on Pricing and the Cost of Electronic Journals, *Journal of Economic Perspectives*, 16, 227-230.
- Dryburgh Associates Ltd (2002), *The Costs of Learned Journal and Book Publishing, A Benchmarking Study for ALPSP*.
- EC Information Society. i2010: Digital Libraries. Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions.
URL: http://europa.eu.int/information_society-activities/digital_libraries/index_en.htm (2005: 9: 30)
- EPIC (2004), Swan, A., P. Needham, S. Proberts, A. Muir, A. O'Brien, C. Oppenheim, R. Hardy and F. Rowland, *Delivery, Management and Access Model for E-prints and Open Access Journals within Further and Higher Education*, A joint report by The Electronic Publishing Innovation Centre (EPIC) in partnership with Key Perspectives Limited.
URL: http://www.jisc.ac.uk/uploaded_documents/ACF1E88.pdf
- Edlin, A. S. and D. L. Rubinfeld (2004), Exclusion or Efficient Pricing? The "Big Deal" Pricing of Academic Journals, *ABA: Antitrust LJ*, 72(1), 128-159.
- Fuegi, D. and M. Jennings (2004) *Updated Version of LIBECON: Library Economics in Europe*, Report for the European Commission, DG Information Society.
- Gatten, J. N. and T. Sanville (2004), An Orderly Retreat from the Big Deal, Is it Possible for Consortia?, *D-Lib Magazine*, 10(10).
- Guédon, J.-C. (2004), The green and gold road to Open-Access: The case for mixing and matching, *Serials Review*, 30(4), 315-327.
- Harnad, S. and Brody, T. (2004), Comparing the Impact of Open Access (OA) vs. Non-OA Articles in the Same Journals, *D-Lib Magazine*, Volume 10 Number 6.
URL: <http://www.dlib.org/dlib/june04/harnad-o6harnad.html>
- Hauge, J.H. (2005), Norway is opening up to Open Access, *ScieCom info*.
URL: http://www.sciecom.org/sciecominfo/artiklar/hauge_05_01.shtml (2005:1)
- Heery, R. and S. Anderson (2005), *Digital Repositories Review*, UKOLN.
- Houghton, J. (2002), The Crisis in Scholarly Communication: An Economic Analysis, Paper presented at the meeting of the Victoria Association for Library Automation 2002.
- House of Commons (2004), The Science and Technology Committee, *Scientific Publications: Free for All*
URL: <http://www.publications.parliament.uk/pa/cm200304/cmselect/cmsstech/399/39902.htm>
- Jeon D.-S. and D. Menicucci (2005), Bundling Electronic Journals and Competition among Publishers, forthcoming in *Journal of the European Economic Association*.
- John Cox Associates Ltd. (2004), *Aggregators and the Primary Journal*, West Sussex, UK: The Association of Learned and Professional Society Publishers.
- Jones, M. (2003), *Archiving E-Journals Consultancy - Final Report*.
URL: http://www.jisc.ac.uk/uploaded_documents/ejournalsfinal.pdf (2003: 10)
- Kyriolidou, M. and M. Young (2004), *ARL Statistics 2002-03*, Washington, DC: Association of Research Libraries.
- Mabe, M. and M. Amin (2001), Growth Dynamic of Scholarly and Scientific Journals, *Scientometrics*, 51(1), 147-162.

- Mabe, M. (2004), CAVEAT AUCTOR: Let the author beware!, *Serials*, 17(1), 41-44.
- Marks, R. H. (1995), The Economic Challenge of Publishing Electronic Journals, *Serials Review: Economics of Electronic Publishing*, 21(1), 85-88.
- McCabe, M. J. (2002), Journal Pricing and Mergers: A Portfolio Approach, *The American Economic Review*, 92(1), 259-269.
- McCabe, M. J. (2004), Information Goods and Endogenous Pricing Strategies: the Case of Academic Journals, *Economics Bulletin*, 12(10), 1-11.
- McCabe, M. J. and C. M. Snyder (2004), The Economics of Open-Access Journals, mimeo, Georgia Institute of Technology.
- Mc Kiernan, G. (2003) Open Archives initiative service providers, *Library Hi Tech News*, Part I: science and technology; Part II: social sciences and humanities; Part III: general.
URL: <http://www.public.iastate.edu/~gerrymck>
- McVeigh, M.E. (2004), *Open Access Journals in the ISI Citation Databases : Analysis of Impact Factors and Citation Patterns. A citation study from Thomson Scientific.*
URL: <http://www.isinet.com/media/presentation/essayspdf/openaccesscitations2.pdf> (2004:10)
- Ministry of Education Finland (2005), *Avoimen tieteellisen julkaisuutoiminnan työryhmän muistio. Opetusministeriön työryhmämuistioita ja selvityksiä*, Report of the Open Access Scientific Publishing Committee.
URL: <http://www.minedu.fi/julkaisut/tiede/2005/tro8/tro8.pdf> (2005:8)
- Minon, M. (2001), *Etude économique et juridique d'un portail pour les revues françaises en sciences humaines et sociales.*
- Minon, M. and G. Chartron (2005), *État des lieux comparatif de l'offre de revues SHS France – Espagne – Italie*, Etude réalisée pour le Ministère de l'éducation nationale de l'enseignement supérieur et de la recherche.
- Morris, S. (2005), The True Costs of Scholarly Journal Publishing, *Learned Publishing*, 18(2), 115-126.
- Morris S. and A. Powell (2005), Dissemination of and access to UK research outputs: response from the Association of Learned and Professional Society Publishers (ALPSP) to the RCUK position paper.
URL: <http://www.alpssp.org/events/2005/RCUK-meeting/RCUKResponse.pdf> (2005:4)
- Muir, A.(2004), *Digital Preservation: awareness, Responsibility and Rights Issues, dans Journal of Information Science,*
URL: <http://www.lboro.ac.uk/departments/ls/disresearch/CLDP/> (2004 :2)
- National Research Council (2004), *Seeking Security: Pathogens, Open Access, and Genome Databases.*
URL: <http://www.nap.edu/catalog/11087.html> (2004:9)
- Nicholas, D., P. Huntigton, B. Russell, A. Watkinson, and C. Tenopir (2005), The Big Deal- Ten Years On, forthcoming in *Learned Publishing*, 18(4).
- Noll, R. and Steinmueller, W. E. (1992), An Economic Analysis of Scientific Journal Prices: Preliminary Results, *Serials Review*, 18(1-2), 32-37.
- Obe, J. S., C. Creaser, and D. Fuegi (2000), *LIBECON: Library Economics in Europe*, Report for the European Commission, DG Information Society.
- OCLC/RLG (2002), *Preservation metadata and the OAIS Information Model. A metadata framework to support the preservation of digital objects.*
URL: http://www.oclc.org/research/projects/pmwg/pm_framework.pdf (2002:6)
- OECD (2004), *Science, Technology and Innovation for the 21st Century. Meeting of the OECD Committee for Scientific and Technological Policy at Ministerial Level, 29–30 January 2004—Final Communiqué.*
URL: http://www.oecd.org/document/0,2340,-en_2649_34487_25998799_1_1_1_1,00.html (2004:1)

- OECD (2005) OECD Working Party on the Information Economy, *Digital Broadband Content: Scientific Publishing*.
URL: www.oecd.org/dataoecd/42/12/35393145.pdf (2005:9)
- Ordoover, J. A. and Willig, R. D. (1978), On the Optimal Provision of Journals qua Sometimes Shared Goods, *American Economic Review*, 68(3), 324-338.
- Oskamp, L.(2003), *Report on current practices among publishers regarding the deposit of digital publications*.
URL: http://www.europeanlibrary.org/pdf/tel_d1.2v1.o.pdf (2003:6)
- PALS (2004), Pathfinder Research on Web-based Repositories Final Report, realised by Mark Ware Consulting, Publisher and Library/Learning Solutions (PALS).
- PREMIS Working Group (2004), *Implementing Preservation Repositories For Digital Materials: Current Practice And Emerging Trends In The Cultural Heritage Community*.
URL: <http://www.oclc.org/research/projects/-pmwg/surveyreport.pdf> (2004:9)
- RCUK (2005) *Position Statement on Access to Research Outputs*.
URL: <http://www.rcuk.ac.uk/access/state-ment.pdf> (2005:6)
- Reaney, D. (2004). *Scholarly Communication: the next 6 years (2005-2010)*, ALPSP.
- Reed Elsevier (2005), Submission to DG Research study on scientific publishing markets in Europe.
- Regazzi, J. (2004), The Shifting Sands of Open Access Publishing: A Publisher's View, *Serials Review*, 30(4), 275-280.
- Research Libraries Group (2002), *Trusted Digital Repositories: Attributes and Responsibilities*, RLG-OCLC.
URL: <http://www.rlg.org/longterm/repositories.pdf>. (2002:5)
- Rightscom Ltd (2005), *Business Model for Journal Content*, commissioned by The Joint Information Systems Committee (JISC).
- Rowlands, I. and D. Nicholas, (2005) *New Journal Publishing Models: an international survey of senior researchers*, CIBER report for the Publishers Association and the International Association of STM Publishers.
- Russell, K. (1999), *Digital preservation: ensuring access to digital materials into the future*.
URL: <http://www.leeds.ac.uk/cedars/Chapter.htm>
- Simba (2004), *Global STM Market Analysis & Forecast 2004*, Stamford, UK: Simba Information.
- Sowden, P. (2004), *University Library Spending on Books, Journals and Electronic Resources*, A report for the Council of Academic and Professional Publishers, London, UK: The Publishers Association.
- Sparks, S. (2005) JISC Disciplinary Differences Report, Rightscom Ltd.
URL: http://www.jisc.ac.uk/index.cfm?name=schol_comms_reports (2005:8)
- SQW Ltd (2003), *Economic Analysis of Scientific Research Publishing*, A report commissioned by the Wellcome Trust, Cambridgeshire, UK: SQW.
- SQW Ltd (2004), *Costs and Business Models in Scientific Research Publishing*, A report commissioned by the Wellcome Trust, Cambridgeshire, UK: SQW.
- Suber (2005), SPARC Open Access Newsletter, issue #86, June 2, 2005.
URL: <http://www.earlham.edu/~peters/fos/-newsletter/06-02-05.htm>
- Susman, T. M. and D. J. Carter (2003), *Publisher Mergers: A Consumer-Based Approach to Antitrust Analysis*, Washington DC: Ropes & Gray LLP and the Information Access Alliance.
- Swan, A. (2005) *American Scientist Open Access Forum*.
URL: <http://www.ecs.soton.ac.uk/~harnad/-Hypermail/Amsci/4337.html> (2005:2)

- Swan, A. and S. Brown (2005), *Open access self-archiving: An author study*, Key Perspectives Limited
URL: [http://www.keyperspectives.co.uk/openaccessarchive/reports/Open%20Access%20II%20\(author%20survey%20on%20self%20archiving\)%202005.pdf](http://www.keyperspectives.co.uk/openaccessarchive/reports/Open%20Access%20II%20(author%20survey%20on%20self%20archiving)%202005.pdf)
- Taylor-Roe, J. and C. Spencer (2005), A Librarian View of Usage Metrics: Through a Glass Darkly, *Serials*, 18(2), 124-131.
- Tenopir, C. (2005), *The role of publishing in the research process*, JISC International Colloquium, International Solutions for the Dissemination of Research: Considering international solutions for purchasing electronic journals and disseminating research outputs, London.
URL: http://www.jisc.ac.uk/index.cfm?name=event_international_o605 (2005:6)
- Tenopir, C. and D. W. King (2000), *Towards Electronic Journals*, Washington DC: SLA Publishing.
- Terry, R. (2005), Funding the way to Open Access, *PLoS Biology*, 3(3), e97.
- Testa, J. and M. E. McVeigh (2004), *The Impact of Open Access Journals. A Citation Study from Thomson ISI*.
URL: <http://www.isinet.com/media/presentrep/acropdf/impact-oa-journals.pdf>
- Watson, J. (2005), You get what you pay for? Archival access to electronic journals. *Serials Review*, article in press.
Doi: 10.1016/j.serrev.2005.06.001
- White, S. and C. Creaser (2004), Scholarly Journal Prices: Selected Trends and Comparisons, *LISU occasional papers no. 34*, commissioned by Oxford University Press.
- Winkel Schwarz, A. (2005), *Denmark's Electronic Research Library: e-publishing initiatives*, ScieCom info.
URL: http://www.sciecom.org/sciecominfo/artiklar/schwarz_04_4.shtml (2004:4)
- WSIS (2003), *Building the Information Society: a global challenge in the new Millennium*.
URL: <http://www.itu.int/wsis/docs/geneva/official/dop.html>

12. LIST OF ACRONYMS

Allea	ALL European Academies
ALPSP	Association of Learned and Professional Society Publishers
CERN	European Organization for Nuclear Research
CNRS	Centre National de la Recherche Scientifique, France
CODATA	Committee on Data for Science and Technology of the International Council for Science
DFG	Deutsche Forschungsgemeinschaft (German Research Foundation)
DG-Research	Directorate General for Research of the European Commission
DOI	Digital Object Identifier
ECARES	European Centre for Advanced Research in Economics and Statistics of the Université Libre de Bruxelles
ESF	European Science Foundation
EUA	European University Association
EUROHORCS	European Heads Of Research Councils
FP	For profit
ICOLC	International Coalition of Library Consortia
ICSTI	international council for scientific and technical information
ICSU	International Council for Science
ICT	Information and Communication Technologies
IDEI	Institut d'Economie Industrielle of the Université des Sciences Sociales de Toulouse
IEEE	Institute of Electrical and Electronics Engineers, Inc.
IMF	International Monetary Fund
INASP	International Network for the Availability of Scientific Publications
INRA	Institut national de la recherche agronomique, France
INSERM	Institut national de la santé et de la recherche médicale, France
IPR	Intellectual Property Rights
ISI	Thomson ISI, newly named Thomson Scientific
JCR	Journal Citation Reports, published by Thomson ISI
JISC	Joint Information Systems Committee, UK
JSTOR	Journal STORage, the Scholarly Journal Archive
LOCKSS	Lots of Copies Keep Stuff Safe
NFP	Not for profit
NIH	National Institutes of Health, United States
NLM	National Library of Medicine, US
OA	Open Access
OAI-PMH	Open Archives Initiative Protocol for Metadata Harvesting

OAIS	Open Archival Information System
OCLC	Online Computer Library Center, Inc.
OECD	Organisation for Economic Co-operation and Development
PLoS	Public Library of Science
PMC	PubMed Central, the U.S. National Institutes of Health free digital archive of biomedical and life sciences journal literature
RCUK	UK Research Councils
RLG	The Research Libraries Group, Inc.
SIAM	Society for Industrial and Applied Mathematics
SPARC	Scholarly Publishing and Academic Resources Coalition
SSH	Social Science and Humanities
STM	Science, Technology and Medicine
TWAS	the academy of sciences for the developing world (originally named "Third World Academy of Sciences")
ULB	Université libre de Bruxelles
UT1	Université des Sciences Sociales de Toulouse
WSIS	World Summit on the Information Society
XML	eXtensible Markup Language

ANNEX TO THE
STUDY ON THE ECONOMIC AND TECHNICAL
EVOLUTION OF THE SCIENTIFIC
PUBLICATION MARKETS IN EUROPE

Information Resources and Contacts

December 2005

Contents

Contents

1. COLLECTION OF DATA AND INFORMATION	3
2. CONTACTS AND MEETINGS WITH EXTERNAL ORGANISATIONS	16
3. PARTICIPATION TO CONFERENCES AND WORKSHOPS	24
APPENDIX 1 : ORGANISATIONS/PERSONS WHO HAVE CONTACTED THE CONTRACTOR	31
APPENDIX 2: PARTICIPANTS AND PERSONS INVITED TO THE “CONSULTATION DAYS”	48

INFORMATION RESOURCES AND CONTACTS

In terms of *information resources and contacts*, the study first builds on a voluminous existing literature. It therefore updates the “state of the art” in terms of reports, studies, surveys, articles... The study also presents data collected for the purpose of an economic analysis, in particular about the journal prices and journal citations. The study has also benefited from considerable interaction with various actors/stakeholders of the sector, be they policy bodies, corporate associations and interest groups. Discussion meetings took place as well as participation and exchange in scientific conferences and policy forums. Three ‘consultation days’ were also organized, where preliminary results of the study were discussed respectively with publisher representatives, research-funding organizations and library representatives.

1. COLLECTION OF DATA AND INFORMATION

REFERENCES, DOCUMENTS AND REPORTS

References, documents and reports that were relevant for the study were collected and deposited in an intranet Website devoted to the study. The intranet has been regularly updated and completed with references and documents, notably those sent by external organisations and publishers associations that wished to contribute.

(See Appendix 1: Organisations/persons that have contacted the contractor.)

DATA ABOUT JOURNAL PRICES

Data on journal prices have been asked to the two major subscription agents that provide print and electronic journal subscription services on the market:

- EBSCO (France) freely provided IDEI with data on journal prices from 1994 to 2004.
- SWETS Information Services (Belgium) provided data about journal prices against payment of 1.500 Euros.

DATA ABOUT PUBLISHERS' COSTS AND ELECTRONIC JOURNAL PRESERVATION STRATEGY

A questionnaire was prepared by ULB Ecares and ULB Library to collect data and information about pricing and costs, and about strategies for the long term preservation of electronic journals.

The questionnaire, together with an introductory letter and guidance notes was sent by email on the 22nd February 2005 to :

- ALPSP Association of Learned and Professional Society Publishers, to be forwarded to the member publishers in Europe and outside Europe, whatever their subject area.
- the Federation of European Publishers to be forwarded, through the FEP member associations, to publishers of scholarly journals in Europe, whatever their subject areas.

The contractor received responses from:

American Physical Society; Armand Colin; Austria Academy of Science Press; Beech Tree Publishing; BioMed Central Ltd; Blackwell Publishing; BMJ; Brill Academic Publishers; Geological Society of London; Lemma Bv; Lippincott Williams & Wilkins; Masson ; Presse Universitaire de France; Presses Universitaires du Mirail ; Royal Society of Chemistry; Royal Statistical Society; Taylor and Francis

Due to the small number of responses and to the incompleteness of the responses received, the data could not be exploited for the study.

2. CONTACTS AND MEETINGS WITH EXTERNAL ORGANISATIONS

CONTACTS WITH THE CONTRACTOR

Following the press release by the European Commission (dated 15/06/2004) announcing the launch of this study, many people from various organisations contacted the contractor: financial analysts, library associations, publishers and publishers associations, scholarly societies, research groups, organisations related to IPR and VAT.

Most of them requested information about the study's scope, objectives and schedule, and offered their contribution to provide information, and to contribute to the debate.

All requests have been answered. Contact persons have been registered for possible further investigation and consultation in the course of study.

See Appendix 1: Organisations/persons that have contacted the contractor.

DISCUSSION MEETINGS WITH THE CONTRACTOR IN BRUSSELS

Upon their request, the contractor has met people representing publishers and publishers associations, political bodies and various organisations related to IST:

- 20/09/2004 Fred Friend, JISC Joint Information Systems Committee UK
- 21/09/2004 Sally Morris, Chief Executive, Association of Learned and Professional Society Publishers
- 22/09/2004 Anne Joseph, Director, EU Governmental Affairs, Reed Elsevier
- 2/12/2004 Anne Bergman-Tahon, Federation of European Publishers; Graham Taylor, Director Council of Academic and Professional Publishers, Publishers Association, UK; Monique Vézinet chargée de mission au Syndicat National de l'Édition.
- 16/02/2005 Prue Backway, Assistant Director, Cross Council Policy & Programmes Office of Science & Technology, and Michel Woodman, Head Business Relations Books, Magazines & e-learning, UK Department of Trade and Industry; Stéphane Goldstein, Policy and Support Manager Research Councils UK.
- 22/03/2005 Pieter Bolman, Chief Executive Officer of STM (International association of scientific, technical and medical publishers) and Mark L. Seeley, Vice President and general Counsel, Elsevier.
- 17/04/2005 Meeting with Mayur Amin and David Tempest, Elsevier Science in their Research Office in oxford.
- 20/04/2005 ALPSP delegation of publishers : Sally Morris, Chief Executive ALPSP; René Olivieri, Chief Executive Officer, Blackwell Publishing; Pippa Powell, Assistant Publications Manager, European Respiratory Society Journals Ltd ; Mike Casey, Publisher, Now Publishers ; Anne Joseph, Director EU Governmental Affairs, Reed Elsevier ; David Brown, British Library.

CONSULTATION DAYS – PRESENTATION OF THE PRELIMINARY RECOMMENDATIONS TO THE ACTORS CONCERNED.

Three “consultation days” were organised in Brussels where preliminary results were discussed respectively with

- (i) publishers representatives, on 12/07/2005;

- (ii) research-funding organisations and scholarly and academic associations, on 5/09/2005;
- (iii) library representative, on 6/09/2005.

The discussions provided useful feedback on the study, especially with respect to the participants’ awareness of the situation for access and dissemination of scientific publications, their respective fears and/or interest regarding the opportunities of ICT in changing the conditions of access and dissemination, and their respective role and responsibility in the changing scholarly publishing process. See Appendix 2: List of participants to consultation days.

3. PARTICIPATION TO CONFERENCES AND WORKSHOPS

UNICA SEMINAR ON SCHOLARLY COMMUNICATION "FAIR PUBLISHING AND FAIR READING",

University of Vienna, Austria, 25-26 November 2004.

Attending: J.-P. Devroey, F. Vandooren, M. Dujardin (ULB Library); M. Dewatripont, A. Walckiers (ULB Ecares); M. Ivaldi (UT1 IDEI); MD Heusse (UT1 Library).

Presentation by Marc Ivaldi (IDEI) and Mathias Dewatripont (ULB ECARES) of the work in progress.

ANTITRUST ISSUES IN SCHOLARLY AND LEGAL PUBLISHING.

Georgetown University Law Center, Washington DC, February 11, 2005 (on invitation only).

Attending: M. Dewatripont, V. Ginsburgh, A. Walckiers. (ULB Ecares)

BERLIN 3 OPEN ACCESS: PROGRESS IN IMPLEMENTING THE BERLIN DECLARATION ON OPEN ACCESS TO KNOWLEDGE IN THE SCIENCES AND HUMANITIES.

University of Southampton, UK, Feb 28th - Mar 1st, 2005.

Presentation by F. Vandooren (ULB library) of the objectives and scope of the EU study. <http://www.eprints.org/berlin3/index.html>

UK RESEARCH COMMUNICATIONS FORUM CONFERENCE.

Oxford, UK, 7 March 2005 (on invitation only)
Forum for publishers, funders, libraries and the Government under the auspices of the UK Department of Trade and Industry,
Attending: F. Vandooren (ULB Library)

THE FUTURE OF THE RESEARCH INFORMATION CHAIN - THE ROLE OF PUBLISHERS AND LEARNED SOCIETIES.

The Hungarian Academy of Science, Budapest, 17 - 18 March 2005.
Jointly organised by ALLEA (*All European Academies*), The European Federation of National Academies of Sciences and Humanities, and STM (*The International Association of Scientific, Technical and Medical Publishers*).
Attending and participating in the panel "Needs of the Research Community–Data and Other Information": P. Legros (ULB Ecares)
http://www.allea.org/cfdata/output/news_detail.cfm?news__id=69

COMMUNICATING SCIENTIFIC, TECHNICAL AND MEDICAL KNOWLEDGE 2005-2010.

Royal Library Belgium, Brussels, 19 April 2005.
Jointly organised by the Belgian Royal Library, Association of Learned and Professional Society Publishers (ALPSP), European Association of Information Services (EUSIDIC), IOS Press Scientific Publishing.
Attending and participating in the panel "Evolution of publishing, for quality and growth": P. Legros (ULB Ecares).
<http://www.alpsp.org/events/c190405.htm>

JISC INTERNATIONAL COLLOQUIUM « INTERNATIONAL SOLUTIONS FOR THE DISSEMINATION OF RESEARCH : CONSIDERING INTERNATIONAL SOLUTIONS FOR PURCHASING ELECTRONIC JOURNALS AND DISSEMINATING RESEARCH OUTPUTS ».

Radisson Edwardian Grafton Hotel, London, 21-22 June 2005 (on invitation only)
Attending: F. Vandooren (ULB Library)

LIBER ANNUAL CONFERENCE 2005.

League of European Research Libraries

Groningen, The Netherlands, 5-8 July 2005.
Presentation by F. Vandooren (ULB Library)
"The EC review of scholarly publishing"
<http://www.kb.dk/guests/intl/liber/>

CER 2005 COMMUNICATING EUROPEAN RESEARCH

Brussels, 14-15 November 2005. Conference organised by the European Commission.
Presentation by M. Dewatripont and participation in the panel: Is there a future for scholarly publishing? Session organised with ALPSP (Association of Learned and Professional Society Publishers)

APPENDIX 1 : ORGANISATIONS/PERSONS WHO HAVE CONTACTED THE CONTRACTOR

FINANCIAL ANALYSTS

Credit Suisse First Boston (Simon Mays-Smith)
3 reports provided: Scientific, Technical and Medical Publishing, 6 April 2004; *Scopus-Moving the goalposts*; *STM Publishing. Evolving Threats to STM publishing*. 29 September 2004.

Exane BNP Paribas (Sami Kassab)
Report provided: Professional Publishing-October 2003

LIBRARY ASSOCIATIONS

CIBER - Coordinamento Interuniversitario Basi dati & Editoria in Rete, Italy (Antonio Fantoni)
Provided CIBER statement on Information and Scientific Communication.

EBLIDA - European Bureau of Library, Information and Documentation Associations (María Pía González Pereira)
Provided the EBLIDA Statement on Open Access.

LIBER - Ligue des bibliothèques européennes de recherche (Paul Ayris)

Procurement for Libraries - UK academic and research library consortia (David Ball).

SCONUL - Society of College, National and University Libraries – UK (Toby Bainton)
Provided the submission made to the UK House of Commons Science and Technology Committee's inquiry on scientific publishing.

SPARC Europe - Scholarly Publishing and Academic Resources Coalition (David Prosser)

PUBLISHERS AND PUBLISHER ASSOCIATIONS

ALPSP - Association of Learned and Professional Society Publishers (Sally Morris)
Provided all the reports commissioned by the ALPSP of interest for the European study (about the costs of publishing, what do societies do with their publishing surpluses, publishers' policies and practices in online publishing and peer review, authors and electronic publication, open access journals).

BioMed Central (Jan Velterop, Marianne Josserand)

FEP - Federation of European Publishers (Anne Bergman-Tahon)
Reports provided: The EU publishing industry: an assessment of competitiveness, EU Enterprise DG; Competitiveness of the European Union publishing industries, EU Enterprise DG; Publishing Market Watch : Sector Report 3: The European Magazine and Journal Market ; Report of FEP activities & European events.

mEDRA - multilingual European DOI Registration Agency (Piero Attanasio)
Report provided: *The use of DOI in eContent value chain*

Public Library of Science PLoS – (Helen Doyle, Andy Gass)

Publishers Association (Graham Taylor)
Reports provided: Scholarly Communication in the digital environment: What do authors want?; University Library Spending on Book, Journals and Electronic Resources - 2004 Update; A report for the Council of Professional and Academic Publishers, by Peter Sowden January 2004.

Reed Elsevier (Anne Joseph)

Report provided: "Submission to DG Research study on scientific publishing markets in Europe". February 2005.

LEARNED SOCIETIES

Max-Planck-Society (Dr Georg W. Botz)

Royal Society of Chemistry (Peter Gregory)

ORGANISATIONS CONCERNED WITH INTELLECTUAL PROPERTY RIGHTS, VAT...

Authors' Licensing & Collecting Society Ltd (Jane Carr)

Report provided on the ALCS position towards the protection of authors' rights

Frankfurt Group (Teresa Hackett)

POLITICAL BODIES

UK Department of Trade and Industry (Prue Backway, Michel Woodman)

UK Office of Fair Trading (Mark Lea, Annette Baxter)

UK Parliament Science and Technology Committee (Emily Commander)
Report: "Scientific publications: Free for all?", July 2004.

RESEARCH FUNDING BODIES

CNRS – Centre National pour la Recherche Scientifique, France (Laurent ROMARY)

DFG Deutsche Forschungsgemeinschaft, Germany (Dr. Johannes Fournier)

Research Councils UK (Stéphane Goldstein)

ORGANISATION FOR STI

EuroScience, European association for the promotion of science and technology (Dr. Françoise Praderie)
Summary and presentations of the conference ESOF 2004, special session "Spreading the word: who profits from scientific publications?";

ICSTI International Council for scientific and technical information (Barry Mahon)

Document provided: "Open Access and the future of Scientific publishing, a white paper"

INIST, Institut national pour l'information scientifique et technique, France (Francis André)

JISC - The Joint Information Systems Committee, UK (Frederick J. Friend)

World Summit On the Information Society (WSIS) Civil Society Working Groups Scientific Information (Dr Francis MUGUET)

RESEARCH CENTRES

CERN (Dr. Hans F Hoffmann)

CNRS Unité d'Indicateurs de Politique Scientifique, France (Serge BAUIN)

German Aerospace Center (DLR) - Deutsches Zentrum für Luft- und Raumfahrt e.V.

Project Management Organisation in the DLR (Andrea Köndgen)

Groupe des Ecoles des Télécommunications, France (Jean-Alain Hernandez)

Swedish School of Economics and Business Administration (HANKEN) Finlande

Partner of the Sci-X project (Scientific Information Exchange) and OACS - Open Access Communication for Science (Bo-Christer Björk)

Technical University of Cottbus, Germany (Andreas Degkwitz)

Provided the Executive summary of a project funded by the Deutsche Forschungsgemeinschaft on the development of the scientific information market in Germany.

APPENDIX 2: PARTICIPANTS AND PERSONS INVITED TO THE “CONSULTATION DAYS”

1. PUBLISHERS AND SUBSCRIPTION AGENCY REPRESENTATIVES, 12 JULY 2005, BRUSSELS

ORGANISATION/COMPANY	NAME OF PARTICIPANT
ALPSP Association of Learned and Professional Society Publishers	Sally Morris
Armand Colin	Jean-Louis Soubret
Blackwell Publishing	René Olivieri
EBSCO	Delphine Dufour
Federation of European Publishers	Anne Bergman-Tahon
Institute of Physics Publishing	Desmond Reaney
Jurist- og Økonomforbundets Forlag	Pia Barnholt Kristoffersen
Lippincott Williams & Wilkins (Europe)	Andrew Richardson
Oxford University Press	Martin Richardson
Public Library of Science	Mark Patterson
Publishers Association, Academic and Professional Division (UK)	Anthony Watkinson
Reed Elsevier	Anne Joseph
Royal Society of Chemistry	Dr Robert J Parker
STM Association	Peter Bolman
Swets Information Services	Bart Vancoppenolle, Yves Van Nieuwenburg
Taylor & Francis	Clare Solomon
Thoams Telford	Leon Heward-Mills
Velterop von Leyden Open Access Consultancy (former Director of BioMed Central)	Jan Velterop

OTHER ORGANISATIONS/COMPANIES WHICH WERE INVITED BUT DID NOT ATTEND

- American Physical Society
- Association des Editeurs Belges
- Association of Subscription Agents and intermediaries
- Austrian Academy of Sciences Press
- Beech Tree Publishing
- BioMed Central Ltd
- BMJ and BMJ Journals
- Brill Academic Publishers
- Casalini
- European Respiratory Society

- Geological Society of London
- German Academic Publishers
- Italian Publishers Association
- Lemma BU
- Masson
- Nature Publishing Group
- Now Publishers
- Presses universitaires de France
- Presses universitaires du Mirail
- Royal Statistical Society
- SNE France
- SPRINGER

2. SCHOLARLY COMMUNITY AND RESEARCH FUNDING BODIES REPRESENTATIVES, 5 SEPTEMBER 2005, BRUSSELS

ORGANISATION	NAME OF PARTICIPANT
Academia Europaea	Prof Arnold Burgen
CERN	Joanne Yeomans
Conférence des Recteurs des Universités Italiennes (CRUI)	Giordana Bruno
Deutsche Forschungsgemeinschaft DFG	Johannes Fournier
European Academies' Science Advisory Council - EASAC	Niceas Schamp
European Life Science Forum ELSF	Willy Stalmans
European Physical Society EPS	Martin C. E. Huber
International Council for scientific and technical Information - ICSTI	Paul Schwander
OTKA Hungria	B. Gábor MAKARA
Swedish School of Economics and Business Administration (HANKEN) Finland	Jonas Holmström
Wellcome Trust	Robert Kiley

OTHER ORGANISATIONS WHICH WERE INVITED BUT DID NOT ATTEND

- Allea All European Academies
- CNR Italy
- CNRS France
- Conférence des Présidents d'Université France
- CSIC Spain
- Czech Science Foundation
- Deutsche Forschungsgemeinschaft DFG
- EASE European Association of Science Editors/EIROforum collaboration between seven European intergovernmental scientific research organisation: CERN, EFDA, EMBL, ESA, ESO, ESRF, ILL
- ESF European Science Foundation
- EUA European University Association
- EUChEMs European Association for Chemical and Molecular Sciences (former ECCC: European Communities Chemistry Council)

- EUROHORCs European Heads of Research Councils
- European Consortium for Political Research (ECPR)
- European Mathematical Society (EMS)
- European Sociological Association (ESA)
- EuroScience
- FNRS Belgium
- FWO Belgium
- ICSU International Council for Science
- ISE (Initiative for Science in Europe)
- Joint Information Systems Committee (UK)
- Max Planck Society
- Netherlands Organisation for Scientific Research (NWO)
- Research Councils UK
- SURF
- Swedish Research Council
- UNICA Network of the Universities of the Capitals of Europe

3. LIBRARY AND LIBRARY CONSORTIA REPRESENTATIVES, 6 SEPTEMBER 2005, BRUSSELS

PARTICIPANTS

ORGANISATION/COMPANY	NAME OF PARTICIPANT
A b-on Biblioteca do Conhecimento Online (b-on) - Portugal	Jose Fernandes
Bibliothèque Interuniversitaire de la Communauté française de Belgique (BICfB) - Belgium	Claudine Kellinckx
Consortium of University Research Libraries (CURL) - UK	Tony Kidd
Consortium universitaire de périodiques numériques (COUPERIN) - France	François Cavalier
Coordinamento Interuniversitario Basi dati & Editoria in Rete (CIBER) - Italy	Eduardo Lasser
Denmark's Electronic Research Library (DEF) - Denmark	Hanne Marie Kværndrup
European Association for Health Information and Libraries (EAHIL)	Suzanne Bakker
Institut de l'Information Scientifique et Technique - Centre National de la Recherche Scientifique (INIST-CNRS) - France	Monique Legentil
Istituto Superiore di Sanità (the Italian National Institute of Health) - Italy	Paola De Castro, Elisabetta Poltronieri
LIBER Ligue des Bibliothèques Européennes de Recherche	Professor Elmar Mittler
OSI Open Access Advocate and JISC Consultant - UK	Fred Friend
Scholarly Publishing and Academic Resources Coalition SPARC Europe	David Prosser
Technical University of Cottbus - Germany	Andreas Degkwitz
The National Electronic Library (FinELib) - Finland	Kristiina Hormia-Poutanen
UKB: Dutch Association of University Libraries, the Royal Library and the Library of the Royal Dutch Academy of Science - The Netherlands	Nol Verhagen

OTHER ORGANISATIONS WHICH WERE INVITED BUT DID NOT ATTEND

- Association of Research Librarians - USA
- Charles University in Prague Library - Czech Republic
- Consorci de Biblioteques Universitaries de Catalunya (CBUC) - Spain
- Consortium of Research Libraries BIBSAM - Sweden
- Consortium of Swiss Academic Libraries
- Consortium Rhine-Westphalia - Germany
- EBLIDA
- Electronic Information for Librarians (eIFL.net)
- ELNET – Estonia
- Eötvös Loránd University of Budapest Library - Hungary
- European Commission on Preservation and Access
- Frankfurt Group
- Friedrich-Althoff-Konsortium - Germany
- HEAL-Link Greece
- He-BIS Konsortium - Germany
- International Coalition of Library Consortia - ICOLC
- Konsortien in Österreich - Austria
- LISU - Loughborough University - UK
- Lithuanian Research Library Consortium
- Poznan Foundation of Scientific Libraries - Poland
- Scottish Confederation of University and Research Libraries (SCURL)
- Society of College, National and University Libraries (SCONUL) - UK
- UK Research Library Network
- University College Dublin Library - Ireland
- University College London Library - UK
- University of Cyprus Library - Cyprus
- University of Latvia Library - Latvia
- University of Zagreb Library - Croatia
- Warsaw University Library - Poland