ECONOMIC ANALYSIS & POLICY, VOL. 39 NO. 1, MARCH 2009

Estimating the Potential Impacts of Open Access to Research Findings

John Houghton and Peter Sheehan
Centre for Strategic Economic Studies
Victoria University
PO Box 14428
Melbourne, Victoria 8001
(email: John.Houghton@vu.edu.au)

Abstract: Advances in information and communication technologies are disrupting traditional models of scholarly publishing, radically changing our capacity to reproduce, distribute, control, and publish information. The key question is whether there are new opportunities and new models for scholarly publishing that would better serve researchers and better communicate and disseminate research findings. Identifying access and efficiency limitations under the subscription publishing model, this paper explores the potential impacts of enhanced access to research outputs using a modified Solow-Swan model, which introduces 'accessibility' and 'efficiency' parameters into calculating returns to R&D. Indicative impact ranges are presented for Government, Higher Education and Australian Research Council R&D expenditures. We conclude that there may be substantial benefits to be gained from more open access to research findings.

I. INTRODUCTION

The existing system of scholarly publishing evolved over many years to serve the needs of disciplinary research in specialist institutions in a print-based environment. The scholarly information environment is now undergoing profound change, with new technologies offering new communication opportunities, changing research practices demanding new capabilities, and an increased focus on research performance (Houghton *et al*. 2003, 2004, Van de Sompel *et al*. 2004, Houghton 2005a, 2005b). As a result, the existing publishing system no longer serves well the needs of researchers for uninhibited access to the research findings of others, the needs of their funders for dissemination of research and for visibility, or the needs of many potential industry and community users for access to research results. A key question facing us today is: are there new opportunities and new models for scholarly publishing that could enhance the dissemination of research findings and, thereby, increase the social returns to investment in R&D?

This paper begins with a brief look at the evolution of scientific publishing during the transition from print to online delivery. It then explores the literature on online access opportunities, focusing on two key aspects: (*i*) access constraints under the subscription publishing system, and (*ii*) the potential impacts of enhanced access on the efficiency and use of R&D. The subsequent section explores the potential impacts of enhanced access to research findings on users in research, industry, government and the wider community. We then modify a simple Solow-Swan model by introducing 'accessibility' and 'efficiency' variables, and use it to estimate the potential impacts of enhanced access on returns to R&D. Results are presented in *Table 1*, which shows plausible ranges of recurring annual gain from a given percentage change in both accessibility and efficiency for Government Expenditure on R&D (GovERD), Higher Education Expenditure on R&D (HERD) and Australian Research Council funded research.

II. THE EVOLUTION OF SCHOLARLY COMMUNICATION

The scholarly publishing and communication system is evolving. Major recent and emerging models for scholarly communication include:

- The 'Big Deal' – where institutional subscribers pay for access to online aggregations of titles through consortial or site licensing arrangements (subscription access is also common for research databases);
- Open Access or 'author-pays' publishing – where authors, their employing or funding organisations contribute to the costs of publication making access free to the reader; and
- Open Access archives and repositories – where organisations support institutional repositories and/or subject archives and authors deposit their work making it freely available to anyone with internet access (which, of itself, does not constitute formal publication).

There are also a number of hybrids, such as delayed open access (*i.e.* allowing open access after a period during which access is restricted to subscribers only), open choice publishing (where authors can choose to pay author fees and make their works open access, or not to pay and make their works subscription only), and less widespread alternatives, such as pay-per-view (Houghton, 2005b).

III. ENHANCED ACCESS OPPORTUNITIES

There is evidence to suggest that toll access publishing has created access limitations, even for researchers in higher education and specialist research centres in developed countries. For example, in a survey of more than 5,500 senior researchers, Rowlands and Nicholas (2005, p. 23) found that almost 74% thought that 'high prices made it difficult to access the journal literature.' Sparks (2005, p. 26-28) reported that almost half of the 750 researchers she surveyed reported having problems gaining access to the resources they needed for their research – with more than half in medical and biological sciences (52.5%) and arts and humanities (53.4%) reporting difficulties. The major problem reported was access to journal articles, books and conference proceedings. Of those reporting difficulties, between 80% and 90% of researchers in medical and biological sciences, physical sciences and social sciences said that their 'library did not take the journals they needed to access for their work', as did

70% to 80% of those in languages, arts and humanities. Looking at discovery services (Jubb *et al*. 2007) noted that:

> 'The main frustration expressed by researchers is… with the problem of then gaining access to the sources and materials they have identified… Their most frequently expressed difficulty is their inability to gain access to journal articles because of a subscription barrier.'

Such studies are complemented by those outlining the potential benefits of enhanced access. An increasing number of studies show that open access articles are used more, both in terms of citations and downloads (HCSTC 2004, Lawrence 2001, Odlyzko 2002, Prosser 2004, Kurtz 2004, Walker 2004, McVeigh 2004, Brody and Harnad 2004, Harnad *et al*. 2004, Brody *et al*. 2005, Getz 2005, Hajjem *et al*. 2005, Davis and Fromerth 2006, etc.). For example, Stevens-Rayburn (2003) noted that Astrophysical Journal articles that are also on the pre-print server have a citation rate twice that of papers not on the pre-print server; Antelman (2004) found a significant difference in the mean citation rates of open-access articles and those that are not freely available online, with the relative increase in citations for open-access articles ranging from a low of 45% in philosophy and 51% in electronic and electrical engineering, to 86% in political science and 91% in mathematics; and Harnad and Brody (2004) noted a study of physics articles published each year between 1992 and 2001 revealing a variation on an annual basis of between 2.5 to 5.8 times more citations for open access articles compared to closed access articles.[1]

A number of authors have pointed to the particular benefits of open access for developing countries, where access to the subscription-based literature has often been limited. Chan et al. (2005, p. 2) noted that: according to a recent survey conducted by the World Health Organization, in the 75 countries with an annual GNP per capita of less then USD 1,000, some 56% of medical institutions had no subscriptions to journals over the last five years; and in countries with a GNP of USD 1,000 to 3,000, 34% had no subscriptions and a further 34% had an average of two subscriptions per year. Conversely, access statistics from OA repositories suggest that researchers from developing countries do use them. For example, during 2007 the ARROW Discovery Service (Australian Research Repositories Online to the World) received hits from some 123 top level domains (i.e. ccTLDs and gTLDs), including, at the lower end of the scale: 1 from the Cook Islands, 2 from Micronesia and Benin, 4 from Tuvalu, 16 from Zimbabwe, Oman, Cambodia, Uganda and Bangladesh, 17 from the Faeroe Islands, East Timor, Georgia and the Virgin Islands, 19 from Brunei Darussalam, 20 from Kazakhstan, 21 from Zambia, 22 from Jordan, Uzbekistan and Qatar, 23 from Namibia, 67 from Papua New Guinea, and so on.[2] A similarly broad range of access is revealed in other repository statistics.

Exploring the advantages of open access institutional repositories, Pinfield (2005) noted the potential for greater research impact, the development of innovative overlay services and new forms of analysis. Looking beyond the research community, Getz (2005, p. 11-12) noted three important dimensions of benefit: broader industry, government and society impacts;

[1] A growing list of such studies is reported by The Open Citation Project (http://opcit.eprints.org/oacitation-biblio.html).

[2] See http://www.arrow.edu.au/

educational impacts; and the potential for greater integration of publications and the other digital objects that are increasingly the outputs of research (*e.g*. numeric data sets, software algorithms, animations, sound and video files). He reported a sevenfold increase in use of the MedLine Index following its move to open access, and 30% use by non-professionals, which clearly suggests that there can be significant impact beyond traditional subscription users.

Kircz (2005) explored the 'dis-benefits' of the subscription publishing system, noting that the published literature was not, as often described, the record of science – at least, not the full record. Firstly, because of timing, it is 'the full stop after the fact' with current discussion in many fields already based on pre-prints and other communications mechanisms (*e.g*. discussion lists, web logs, etc.). Secondly, because of selectivity in publishing, it is 'only a trophy cabinet' with little reporting in the formal journal literature of failed experiments, and trial and error tests, etc. The latter was also noted by Gallagher (2005, p. 8), who suggested that repositories would be 'more likely than existing journals to include accessible archives of negative data.' These points highlight two important advantages of open access for the efficiency of R&D: (*i*) timeliness and speed of reporting, especially through the posting of pre-prints; and (*ii*) the potential to create a fuller record of science through mandated deposit of findings and other not previously reported materials (*e.g*. field notes or laboratory notes, related data sets, etc.), thus speeding up the research process and avoiding the inefficiency of duplicative research and the pursuit of blind alleys.

IV. IDENTIFYING THE IMPACTS THAT MIGHT BE MEASURED

Publication is an important mechanism for accelerating the rate of technological innovation (Sorenson and Fleming 2004). The potentially measurable impacts of enhanced access to research findings (research data and publications) relate to their use by other researchers, industry and government use, and potential use by individuals in the wider community.

4.1. Research

The most immediate impacts of enhanced access would be likely to be felt within research, wherein the dimensions of potential impact include:

- Faster access, speeding up the research and discovery process, increasing returns to investment in R&D and, potentially, reducing the time/cost involved for a given outcome and improving the efficiency of R&D;
- Improved access leading to better informed research, reducing the pursuit of blind alleys and reducing duplicative research, saving wasted and duplicative R&D expenditure and improving the efficiency of R&D;
- Wider access both providing enhanced opportunities for multi-disciplinary research, inter-institutional and inter-sectoral collaborations, and enabling researchers to study their context more broadly, potentially leading to increased opportunities for and rates of commercialisation; and
- Greater educational access leading to improved education outcomes, enabling a given education spend to produce a higher level of educational attainment, leading to enhancement of the capabilities of future researchers and research users.

More open access to the findings of publicly funded research might also usefully serve to counteract the tendency for limited disclosure and the, so called, secrecy problem (Van Looy *et al*. 2006) or privatisation of scientific commons (Nelson 2004).

4.2. Industry and Government

Given relative levels of access under the subscription publishing system, and recognising that industry gets most of its benefit from academic science through open channels (Cohen *et al*. 2002, Nelson 2004), it is possible that greater potential impacts lie in enhanced disclosure and greater access for industry and government users. The dimensions of potential impact include:

- The potential for wider access to both accelerate and widen opportunities for adoption and commercialisation of research findings, thereby increasing returns on public investment in R&D and on private investment in discovery and commercialisation;
- The potential for much wider access than the subscription publishing system gives for doctors/nurses, teachers/students, small firms in consulting, engineering, architecture, design, electronics, ICT, biotechnology, nanotechnology, etc., with a positive impact on quality of service and, possibly, productivity in both those sectors of the economy and those of their customers and clients; and
- The potential for the emergence of new industries based upon open access content – there are examples of new industries built on publicly accessible data (*e.g*. weather derivatives based on meteorological data (Stell, 2005)), and there are potential futures for publishers to become value adding services providers overlaying open access content (*e.g*. providing discovery and access tools, peer review services, bibliometrics and webometrics for research evaluation, etc.), which might, in turn, increase the efficiency of R&D, improve quality and enhance research evaluation and lead to better focused R&D expenditures.

Impacts might be felt more in particular sectors (*e.g*. knowledge intensive services, biotechnology, etc.). Impacts in such areas as management, economic and engineering consulting might be significant, raising the quality of advice to the benefit of customers and clients across the economy. There may also be positive impacts on policy development, through better informed policy debate and enhanced access to the information underpinning policy decisions. One particularly important dimension might be the potential for greater access for small and medium sized firms (SMEs), enabling small firms to do more research internally and increasing their ability to absorb research from outside, and to innovate.

4.3. The Wider Community

In relation to the wider community, the dimensions of potential impact include the potential contribution of open access to the 'informed citizen' and 'informed consumer' – with implications for better use of health and education services, better consumption choices, etc. leading to greater welfare benefits, better health and education outcomes, etc., which may in turn lead to productivity improvements. There is also the issue of the role of knowledge in the functioning of a market economy (Kamperman Sanders 2006, p. 856).

V. QUANTIFYING THE IMPACTS OF ENHANCED ACCESS

The task of fully exploring the impacts of enhanced access is substantial, but it is possible to gain some sense of the scale of potential impacts by modifying the Solow-Swan model, introducing '*accessibility*' and '*efficiency*' parameters into calculating the returns to R&D. While there are recognised limitations to the traditional approach to estimating returns to R&D (Salter and Martin 2001, Scott *et al*. 2002, Shanks and Zheng 2006, Martin and Tang 2007), it does provide one basis for preliminary estimates of the potential impacts of enhanced access.

5.1. Returns to R&D in a Simple Growth Model

In the basic Solow-Swan model, the key elements are a production function:

$$Y = A^{\eta} K^{\beta} L^{\alpha} \tag{1}$$

where A is an index of technology, K is the capital stock and L is the supply of labour, with both K and L are taken to be fully employed by virtue of the competitive markets assumption, and an accumulation equation:

$$\dot{K} = sY - \delta K, \tag{2}$$

where $\dot{K}$ is the net investment or the change in the net capital stock, equal to gross investment less depreciation, and δ is a constant depreciation rate. Substituting (1) into (2) gives

$$= sA^{\eta} K^{\beta} L^{\alpha} - \delta K. \tag{3}$$

From (3) it is possible to determine the conditions for steady state growth in the capital stock.

Re-arranging, taking logarithms, differentiating with respect to time and imposing the condition that for steady state growth:

$$d/dt(\ln \dot{K}/K) = 0$$

gives:

$$\dot{K}/K = \frac{\eta}{1-\beta} \dot{A}/A + \frac{\alpha}{1-\beta} \dot{L}/L \tag{4}$$

where $\dot{K}/K = \dot{C}/C = \dot{Y}/Y$, is the single constant steady state rate of growth of capital stock, consumption and output, respectively.

The main features of the Solow-Swan model are apparent from equation (4). Firstly, if technology and labour supply are fixed, the steady state growth rate is zero. That is, there is no endogenous growth in the model, growth being driven in the steady state by change in the exogenous variables. Secondly, if one of technology and population show positive growth then the steady state growth rate of the economy is proportional to the growth rate in that variable; if both rates are positive the economy's growth rate is a weighted average of the two. Thirdly, the steady state growth rate does not depend on either the level of savings or of investment in the economy. An economy that continuously saves and invests 20% of national income will have a higher level of output than one investing 5%, but it will not have a higher steady state growth rate. Thus the broad economic message of the Solow-Swan model is that steady growth is possible in a purely competitive world, provided that there is growth in either population or technology, or both.

1. Contributions to Growth and Total Factor Productivity

Solow (1957) further developed this model in a way that provided the foundations for subsequent 'growth accounting'. Starting with total differentiation of the production function (1), and substituting for the partial derivatives of Y from (1) with respect to each of its arguments, yields:

$$\dot{Y}/Y = \eta\dot{A}/A + \beta\,\dot{K}/K + \alpha\dot{L}/L. \tag{5}$$

Equation (5) can then be used in two main ways in the empirical study of growth.

Given that in the competitive model capital and labour are paid their marginal products and assuming constant returns to scale, β and α can be estimated from the relative shares of capital and labour. A variant of (5) with those weights can then be used to estimate the relative contribution of capital, labour, technology and other factors to growth. Solow made pioneering estimates in 1957, the results of which he later described as 'startling' (Solow 1987), and these have been much refined and amplified by Denison (1985) and others. Solow found that $7/8^{th}$ of the growth in real output per worker in the US economy between 1909 and 1949 was due to 'technical change in the broadest sense' and only $1/8^{th}$ to capital formation. Denison's 1985 estimates covered the US economy for the period 1929 to 1982. Of the growth in real business output of 3.1% per annum over that period, he found that the increase in labour input with constant educational qualifications accounted for about 25% and capital input for 12%. Most of the remainder is accounted for by technological progress and by the increased human capital of the work-force. What was 'startling' about these results was the relatively minor contribution to output growth arising from the increase in the traditional factors of production, capital and labour.

The other related use of equation (5) is to estimate the 'Solow residual', or total factor productivity. This is defined as the difference between output growth and the weighted sum of the growth rates of factor inputs (K and L), using constant return to scale weights. That is, total factor productivity growth (TFP) is given by:

$$TFP = \dot{Y}/Y - \beta\dot{K}/K - \alpha\dot{L}/L, \tag{6}$$

where $\beta = 1 - \alpha$, and β and α are derived from the shares of capital and labour in total income.

Total factor productivity is thus the growth in output not accounted for, on these assumptions, by the growth in capital and labour inputs. This method is now used very widely around the world in measuring productivity. This recent use has confirmed the broad Solow-Denison findings, in that for most modern economies total factor productivity growth is significantly more important than expansion of inputs in explaining total output growth. However, it must be remembered that the method rests on the assumptions embedded in the Solow-Swan model and that, as a consequence, the finding that the larger proportion of growth is to be explained by an exogenous 'technical change in the broadest sense' constitutes something of an admission of defeat for economic analysis.

2. Estimating the Rate of Return to R&D

While there are recognised limitations to the traditional growth model approach, this basic framework has been widely used in estimating the rate of return to R&D. The standard approach to estimating returns to R&D is to divide the technology variable A in (1) into two components, a stock of R&D knowledge variable R and a variable Z that represents a matrix of other factors affecting productivity growth. The production function then becomes:

$$Y = K^{\alpha} L^{\beta} R^{\gamma} Z^{\eta}, \qquad (7)$$

and the counterpart of equation (5) becomes:

$$\dot{Y}/Y = \alpha\, \dot{K}/K + \beta\, \dot{L}/L + \gamma\, \dot{R}/R + \eta \dot{Z}/Z. \qquad (8)$$

That is, the rate of growth of the R&D knowledge stock (*i.e.* accumulated R&D expenditure or R&D capital) contributes to output growth as a factor of production, with elasticity γ. The rate of return to knowledge (∂y/∂R) is that continuing average per cent increment in output resulting from a one per cent increase in the knowledge stock. This can be readily derived from the elasticity γ by

$$\partial y/\partial R = \gamma. (Y/R). \qquad (9)$$

The normal approach to creating a measure of the stock of R&D knowledge, for a given industry or for the economy as a whole, is to use the perpetual inventory method to create the knowledge stock from the flows of R&D, using the relationship:

$$R_t = (1 - \delta)\, R_{t-1} + R\&D_{t-1}, \qquad (10)$$

where δ is the rate of obsolescence of the knowledge stock. This method also requires some starting estimates (R_0) of the knowledge stock, and estimates can be sensitive to that assumption.

Then the knowledge stock at time *t* is given by:

$$R_t = (1 - \delta)^t R_0 + \sum_{i=0}^{t-1} (1 - \delta)^I R\&D_{t-1} \qquad (11)$$

Given a series for R and for the variables Z, it is then possible to estimate γ by either of the two methods noted above: estimate equation (8) with the parameters α .. η unconstrained, or obtain estimates of the parameters α and β (constrained to be equal to one) from the factor shares of capital and labour, calculate TFP by a variant of (7) and regress R and Z on TFP to obtain γ.

3. Incorporating the Efficiency of Research and Accessibility of Knowledge

This standard approach makes some key simplifying assumptions. Here we note three in particular. It is assumed that:

- All R&D generates knowledge that is useful in economic or social terms (*efficiency of R&D*);
- All knowledge is equally accessible to all entities that could make productive use of it (*accessibility of knowledge*); and
- All types of knowledge are equally substitutable across firms and uses (*substitutability*).

A good deal of work has been done to address the fact that the substitutability assumption is not realistic, as particular types of knowledge are often specialised to particular industries and applications. Much less has been done on the other two assumptions, which are our focus.

We define an '*accessibility*' parameter ε as the proportion of the R&D knowledge stock that is accessible to those who could use it productively, and an '*efficiency*' of R&D parameter ϕ as the proportion of R&D spending that generates useful knowledge. Then starting with a given stock of useful knowledge R^*_0 at the start of period zero, useful knowledge at the start of period 1 will be given by:

$$R^*_1 = (1 - \delta)\, R^*_0 + \phi\, R\&D_0, \tag{12}$$

where the contribution of R&D in period zero to the knowledge stock is reduced by the parameter ϕ to allow for unproductive R&D. This means that the stock of useful knowledge at period t is given by:

$$R^*_t = (1 - \delta)^t\, R^*_0 + \phi \sum_{i=0}^{t-1} (1 - \delta)^i\, R\&D_{t-1} \tag{13}$$

If the period over which knowledge is accumulated is long, so that $(1 - \delta)^t R^*_0$ is small relative to R^*_t, then R^*_t can be approximated by ϕR. However, only a proportion of useful knowledge may be accessible, so that accessible useful knowledge at period t is εR^*_t, and hence approximately $\phi\varepsilon R_t$, where R_t is the stock of knowledge as calculated under the standard methods.

Using this approximation and noting that it is accessible useful knowledge that is the correct factor in the production function, (6) becomes:

$$Y = K^{\alpha} L^{\beta} (\phi\varepsilon R)^{\gamma} Z^{\eta} \tag{14}$$

If ϕ and ε are independent functions of time, then the results of estimating a linearised version of (14) that excludes them will be misleading. However, if we assume that these parameters reflect institutional structures for research and research commercialisation in a given country, and can hence be taken as fixed (and as less than or equal to one), then the standard results stand, but need to be reinterpreted. Again using R as the stock of knowledge calculated by the standard method (which assumes $\phi = \varepsilon = 1$) and R^* as the corresponding accessible stock of useful knowledge, then $R = R^*/\phi\varepsilon$, and the rate of return to useful and accessible knowledge becomes:

$$\partial y/\partial R^* = \gamma. (Y/R^*) = \gamma/\phi\varepsilon. (Y/ R) = \gamma. (Y/R).1/\phi\varepsilon. \quad (15)$$

Thus, if ϕ and/or ε are less than one, the rate of return to R^* is greater than that to R by the factor $1/\phi\varepsilon$. This does not imply that the measured rate of return to R is biased, because $R^* = \phi\varepsilon R$.

Assume now that there is a one-off increase in the value of ϕ and ε, from the constant values of ϕ_0 and ε_0 to new values of $(1 + \delta_\phi)\phi_0$ and $(1 + \delta_\varepsilon)\varepsilon_0$, respectively. Then the rate of return to R^*, that is:

$$\partial y/\partial R^* = \gamma. (Y/R). (1/\phi_0\varepsilon_0) \quad (16)$$

is fixed, but the return to R will increase:

$$\partial y/\partial R = \gamma. (Y/R) = \phi_1\varepsilon_1 \, \partial y/\partial R^* = \gamma. (Y/R). (\phi_1\varepsilon_1 /\phi_0\varepsilon_0) \quad (17)$$

$$= \gamma. (Y/R). (1 + \delta_\phi).(1 + \delta_\varepsilon)\varepsilon_0.$$

It follows from (17) that, because the increase in efficiency and accessibility leads to a higher value of R^* for a given level of R, the rate of return to R will increase by the compound rate of increase of the percentage changes in ϕ and ε.

The basic result of the foregoing is that, if '*accessibility*' and '*efficiency*' are constant over the estimation period, but then show a one-off increase (*e.g*. because of a move to open access), then, to a close approximation, the return to R&D will increase by the same percentage increase as that in the accessibility and efficiency parameters.

VI. ESTIMATING THE IMPACTS OF A ONE-OFF INCREASE IN ACCESSIBILITY AND EFFICIENCY

Open access discussion focuses on the accessibility of research data and publications from publicly funded research, such as that undertaken in government laboratories and universities. Therefore, by way of demonstration, we have estimated the recurring annual gain from a given percentage change in accessibility and efficiency for Government Expenditure on R&D (GovERD), Higher Education Expenditure on R&D (HERD) and Australian Research Council funded research in Australia circa 2006-07. *Table 1* presents plausible impact ranges based on social rates of return to R&D of 20% to 60% (Arundel and Geuna 2004), and 1% to 10% increases in accessibility and efficiency (Houghton *et al*. 2009). It shows that whether applied across the board or to sector specific research findings (*e.g*. open access to publicly funded or higher education research) there may be substantial potential benefits to be gained from more open access to research findings.

Table 1: Estimates of the impacts of a on-off increase in accessibility and efficiency on returns to R&D, 2006-07 (AUD, millions)

Government expenditure on R&D	Rate of return to R&D				
AUD 8,026 million	20%	30%	40%	50%	60%
Per cent change in accessibility and efficiency	Recurring annual gain from increased accessibility & efficiency (million)				
1%	32	48	65	81	97
2%	65	97	130	162	195
5%	165	247	329	411	494
10%	337	506	674	843	1,011
Higher Education expenditure on R&D	Rate of return to R&D				
AUD 5,404 million	20%	30%	40%	50%	60%
Per cent change in accessibility and efficiency	Recurring annual gain from increased accessibility & efficiency (million)				
1%	22	33	43	54	65
2%	44	66	87	109	131
5%	111	166	222	277	332
10%	227	340	454	567	681
Australian Research Council research funding	Rate of return to R&D				
AUD 572 million	20%	30%	40%	50%	60%
Per cent change in accessibility and efficiency	Recurring annual gain from increased accessibility & efficiency (million)				
1%	2	3	5	6	7
2%	5	7	9	12	14
5%	12	18	23	29	35
10%	24	36	48	60	72

Sources: R&D expenditure data from ABS (2008) Research and Experimental Development, All Sector Summary, Australia, 2006-07, Cat No 8112.0 and ARC (2007) ARC Budget 2006-07, ARC, Canberra.

For the purposes of discussion we take the lower bound 20% as indicative of returns to publicly funded R&D, and assume that the change (*e.g.* to open access) has no net impact on the rates of accumulation or obsolescence of the stock of knowledge. While there are measurement problems and the accessibility and efficiency parameters would be difficult to distinguish if one were to run regressions, they are conceptually different and one could develop proxy indicators to determine a plausible range for their values. For example, one could use the open access citation advantage as a proxy indicator of the potential impacts of enhanced accessibility.

Reading and writing habits vary between locations and fields of research, but ongoing surveys of research activities suggest that, on average, active researchers spend around 20% to 25% of their time reading and writing journal articles (Tenopir and King 2000). Hence, as a conservative estimate we assume that 20% of the stock of research knowledge resulting from research time and expenditure relates to journal articles. There have been many studies examining the potential citation advantage of open access. The observed advantages vary considerably, with Hajjem *et al*. (2005) concluding that:

> In 2001, Lawrence found that articles in computer science that were openly accessible (OA) on the Web were cited substantially more than those that were not. We have since replicated this effect in physics. To further test its cross-disciplinary generality, we used 1,307,038 articles published across 12 years (1992-2003) in 10 disciplines (Biology, Psychology, Sociology, Health, Political Science, Economics, Education, Law, Business, Management). The overall percentage of OA (relative to total OA + NOA) articles varies from 5%-16% (depending on discipline, year and country) and is slowly climbing annually. Comparing OA and NOA articles in the same journal/year, OA articles have consistently more citations, the advantage varying from 25%-250% by discipline and year.

Hence, as starting point, one might take 25% as a conservative estimate of the potential citation advantage from open access. Combining these would suggest that a 5% increase in accessibility might be plausible (*i.e*. a 25% increase in use of 20% of the research stock of knowledge). With many possible impacts on efficiency, but few immediately apparent metrics, we assume a similar 5% increase in efficiency (for illustrative purposes).

Given a 20% social return to publicly funded R&D, for the major categories of R&D expenditure in Australia in 2006-07 a 5% increase in accessibility and efficiency would have been worth:

- AUD 165 million per annum in increased returns to Government Expenditure on R&D, or approximately 0.02% of Gross Domestic Product;
- AUD 111 million per annum in increased returns to Higher Education Expenditure on R&D; and
- AUD 12 million per annum in increased returns to Australian Research Council funded R&D.

These are recurring annual gains from the effect of one year's R&D, so if the change that brings increases in accessibility and efficiency (*e.g*. a shift to open access publishing) is permanent they can be converted to growth rate effects. For example, assuming a 10% discount rate and 5% per annum growth in Government Expenditure on R&D, these increased returns might be worth around AUD 9 billion over 20 years (Net Present Value).

While it is impossible to calculate the quantum of benefits with certainty, these simple estimates of the potential impacts of enhanced access on returns to R&D suggest that a move towards more open access to research findings may have a substantial impact.

VII. CONCLUSIONS AND DEVELOPMENTS

There are two main conclusions from this analysis. First, while there are many limitations, introducing accessibility and efficiency parameters into the standard returns to R&D approach offers one possible foundation for estimating the potential impacts of enhanced/open access.

Second, these preliminary estimates suggest that there may be substantial benefits to be gained from more open access.

There are many ways in which these preliminary estimates might be refined, including, for example:

- Developing more grounded estimates of the contribution of research data, publications and other potentially open access digital objects to the stock of knowledge generated by R&D (*i.e.* estimating what proportion of the stock of knowledge might be affected by open access);
- Finding direct measurable links between access, use and efficiency, to enable estimation of the appropriate percentage change to apply to the accessibility and efficiency variables (*e.g.* direct links between access, downloads and citations, and access and levels of duplicative activity);
- Finding measures of and comparing the use of subscription-based and open access content by non-research users (*e.g.* users in industry, government, non-government organisations and the wider community);
- Exploring the extent of leakage of impacts across national borders, to better understand where benefits may accrue and the potential importance of international initiatives towards open access;
- Exploring the impacts of enhanced access to different R&D categories (*e.g.* Gross Expenditure on R&D, Government Expenditure on R&D, Higher Education R&D Expenditure, etc.) using appropriate rates of returns for each category;
- Further examining the potential impacts of a one-off change to enhanced access on the accumulation and obsolescence of the stock of research knowledge, and exploring the impacts of applying different rates of depreciation; and
- Exploring marginal as well as average rates of return to R&D, and the potential impacts of increases in accessibility and efficiency on marginal returns to R&D.

Given substantial R&D expenditures and the scale of the potential impacts identified in this preliminary work, these issues represent fertile ground for further policy relevant inquiry.

ACKNOWLEDGEMENTS

The authors would like to acknowledge the support of the Systemic Infrastructure Initiative, a part of the Australian Commonwealth Government's *Backing Australia's Ability – An Innovation Action Plan for the Future*, and The UK Joint Information Systems Committee (JISC).

REFERENCES

ABS (2008). *Research and Experimental Development, All Sector Summary, Australia, 2006-07*, Cat No 8112.0, Australian Bureau of Statistics, Canberra.

Antelman, K. (2004). Do open-access articles have a greater research impact?, *College & Research Libraries News*. 65: 372-382.

ARC (2007). *ARC Budget 2006-07*, Australian Research Council, Canberra.

Arundel, A. and A. Geuna (2004). 'Proximity and the use of public science by innovative European firms,' *Economics of Innovation and New Technology.* 3: 559-580.

Brody, T. and S. Harnad (2004). Comparing the Impact of Open Access (OA) vs. Non-OA Articles in the Same Journals, *D-Lib Magazine*. 10(6).

Brody, T., S. Harnad, L. Carr (2005). Earlier Web Usage Statistics as Predictors of Later Citation Impact, *Journal of the American Association for Information Science and Technology (JASIST)*. 57: 1060-1072.

Chan, L., B. Kirsop, B. and S. Arunachalam (2005). Open Access Archiving: the fast track to building research capacity in developing countries, *SciDevNet*. November 2005.

Coe, D.T and E. Helpman (1993). *International R&D Spillovers*. NBER Working Paper 4444, National Bureau of Economic Research, Cambridge MA.

Cohen, W., R. Nelson, and J. Walsh (2002). Links and impacts: the influence of public research on industrial R and D, *Management Science*. 48: 1-23.

Davis, P.M. and M.J. Fromerth, (2006). *Does arXiv lead to higher citations and reduced publisher downloads for mathematics articles?* arXiv.org, 14 March 2006.

Denison, E.F. (1985). *Trends in American Economic Growth, 1929-1982*. Washington DC: Brookings Institution.

Dowrick, S. (2003). *A Review of the Evidence on Science, R&D and Productivity.* Department of Education, Science and Training, Canberra.

Gallager, R. (2005). Why we need institutional repositories, *The Scientist*. 19(19): 8.

Getz, M. (2005). Open Scholarship and Research Universities. Department of Economics Working Paper No. 0517, Vanderbilt University, Nashville.

Godin, B. and C. Doré (2004). Measuring the Impacts of Science: Beyond the Economic Dimension. CSIIC Working Paper, CSIIC, Canada.

Gorman, G.E. and F. Rowland (2005). *International Year Book of Library and Information Management 2004-2005: Scholarly Publishing in an Electronic Era*. London: Facet Publishing.

Griliches, Z. (1995). R&D and productivity: Econometric Results and Measurement Issues. In Stoneman, P. (ed.), 1995. *Handbook of The Economics of Innovation and Technological Change*. Oxford UK: Blackwell, p. 52–89.

Hajjem, C., S. Harnad, S., and Y. Gingras (2005). Ten-Year Cross-Disciplinary Comparison of the Growth of Open Access and How it Increases Research Citation Impact, *IEEE Data Engineering Bulletin*. 28: 39-47.

Harnad, S. and T. Brody (2004). Comparing the impact of open access (OA) vs. non-OA articles in the same journals, *D-Lib Magazine*. 10(6) : June 2004.

Harnad, S., T. Brody, F. Vallieres, L. Carr, S. Hitchcock, G. Yves, O. Charles, H. Stamerjohanns, and E. Hilf (2004). The Access/Impact Problem and the Green and Gold Roads to Open Access, *Serials Review.* 30(4).

HCSTC (House of Commons Science and Technology Committee), 2004. *Scientific Publications*: *Free for all?* Tenth Report of Session 2003-04, The Stationery Office, London.

Houghton, J.W. (2005a). Economics of Publishing and the Future of Scholarly Communication. In: (eds.) G.E. Gorman and F. Rowland, 2005. *International Year Book of Library and Information Management 2004-2005: Scholarly Publishing in an Electronic Era*. London: Facet Publishing, p. 165-188.

Houghton, J.W. (2005b). *Digital Delivery of Content: Scientific Publishing*. Organisation for Economic Cooperation and Development, Paris.

Houghton, J.W. (2005c). Changing Research Practices and Research Infrastructure Development, *Higher Education Management and Policy.* 17: 63-81.

Houghton, J.W., C. Steele, and M. Henty (2003). *Changing Research Practices in the Digital Information and Communication Environment*. Department on Education, Science and Training, Canberra.

Houghton, J.W., C. Steele, and M. Henty (2004). Research Practices, Evaluation and Infrastructure in the Digital Environment, *Australian Academic & Research Libraries*. 35: 161-176.

Houghton, J.W., C. Steele, and P.J. Sheehan (2006). *Research Communication Costs in Australia, Emerging Opportunities and Benefits*. Department of Educations, Science and Training, Canberra.

Houghton, J.W., B. Rasmussen, P.J. Sheehan, C. Oppenheim, A. Morris, C. Creaser, H. Greenwood, M. Summers, and A. Gourlay (2009). *Economic Implications of Alternative Scholarly Publishing Models: Exploring the Costs and Benefits*, Report to The Joint Information Systems Committee (JISC) by Victoria University & Loughborough University. http://ie-repository.jisc.ac.uk/278/3/EI-ASPM_Report.pdf

Industry Commission, 1995. *Research and Development*. Report No 44, Industry Commission, Canberra.

Jones, C.I., and J.C. Williams (1998). Measuring the Social Return to R&D**,** *Quarterly Journal of Economics.* 113: 1119-1135.

Jubb, M., H. Look, and S. Sparks (2007). 'Researchers' use and perceptions of discovery services,' *Learned Publishing*. 20: 147-153.

Kamperman Sanders, A. (2006). Limits to database protection: Fair use and scientific research exemptions, *Research Policy.* 35: 854-874.

Kircz, J. (2005). Institutional Repositories, a new platform in Higher Education and Research. Discussion Paper Commissioned by SURF, KRA Publishing, Amsterdam.

Kurtz, M.J. (2004). Restrictive access policies cut readership of electronic research journals articles by a factor of two. Presented at National Policies on Open Access (OA) Provision for University Research Output, New College, Southampton.

Lawrence, S. (2001). Free online availability substantially increases a paper's impact, *Nature*. May 2001: 411-521.

Martin, B.R. and P. Tang, P. (2007). The benefits of publicly funded research, SWEPS Paper No. 161, Science Policy Research Unit, University of Sussex, Brighton.

McVeigh, M.E. (2004). *Open Access Journals and the ISI Citation Database: Analysis of Impact Factors and Citation Patterns*. Thomson Scientific Whitepaper, Philadelphia: Thomson Scientific.

Nelson, R.R. (2004). The market economy and the scientific commons, *Research Policy.* 33: 455-471.

NLA (2006). *Statistics for ARROW Discovery Service*. National Library of Australia, Canberra.

Odlyzko, A. (2002). The rapid evolution of scholarly communication, *Learned Publishing*. 15: 7-19.

OECD (2006). *Main Science and Technology Indicators 2005*. Organisation for Economic Cooperation and Development, Paris.

Pinfield, S. (2005). A mandate to self archive? The role of open access institutional repositories, *Serials.* 18: 30-34.

Productivity Commission (2006). *Public Support for Science and Innovation: Draft Report*. Productivity Commission, Canberra.

Prosser, D. (2004). Between a rock and a hard place: The big squeeze for small publishers, *Learned Publishing*. 17: 17-22.

Rankin, J. (2005). *Institutional Repositories for the Research Sector: Feasibility Study.* Wellington: National Library of New Zealand.

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, Amsterdam.

Salter, A.J. and B.R. Martin (2001). The economic benefits of publicly funded basic research: a critical review, *Research Policy.* 30: 509-532.

Scott, A., G. Steyn, A. Geuna, S. Brusoni, and E. Steinmueller (2002). *The Economic Returns to Basic Research and the Benefits of University-Industry Relationships*. Report to the Office of Science and Technology, London.

Shanks, S. and S. Zheng (2006). Econometric modelling of R&D and Australia's productivity. Staff Working Paper, Productivity Commission, Canberra.

Solow, R.M. (1957). Technical Change and the Aggregate Production Function, *Review of Economics and Statistics*. 39: 312-320.

Solow, R.M. (1987). Growth Theory and After. *R.M. Solow – Prize Lecture*, Nobel e-Museum Laureates.

Sorenson, O. and L. Flemming (2004). Science and the diffusion of knowledge, *Research Policy.* 33: 1615-1634.

Sparks, S. (2005). *JISC Disciplinary Differences Report*. London: Rightscom.

Stell, J. (2005). *Results of the 2005 PwC Survey.* Presentation to the Weather Risk Management Association, November 2005.

Stevens-Rayburn, S. (2003). *ASTRO: report from the AAS Publications Board meeting*. November 2003.

Stoneman, P. (ed.), 1995. *Handbook of The Economics of Innovation and Technological Change*. Oxford UK: Blackwell.

Tenopir, C. and D.W. King (2000). *Towards Electronic Journals: Realities for Scientists, Librarians and Publishers*. Washington D.C.: Special Libraries Association.

Van de Sompel, H., S. Payette, J. Erickson, C. Lagoze, and S. Warner (2004). Rethinking Scholarly Communication: Building the system that scholars deserve, *D-Lib Magazine*. 10(9).

Van Looy, B., J. Callaert, and K. Debackere (2006). Publication and patent behaviour of academic researchers: Conflicting, reinforcing or merely co-existing?, *Research Policy.* 35: 596-608.

Walker, T.J. (2004). Open Access by the article: An idea whose time has come?, *Nature*. Web Focus: Access to the Literature. June 2004.