MODELING A MODEST PROPOSAL FOR INCREASING THE EFFICIENCY OF ACADEMIC RESEARCH DISSEMINATION

Hugh M. Cannon
Wayne State University
hugh.cannon@wayne.edu

James N. Cannon
Iowa State University
jcannon@iastate.edu

ABSTRACT

Over the years, ABSEL has sought to promote research by making its research more readily available to scholars. This parallels a broader effort by academic research institutions to improve the dissemination of scholarly research in general. To put these efforts in perspective, we view academic journals and conference proceedings as a kind of distribution intermediary for research, analogous to retail intermediaries in conventional marketing. In an environment of increasing competition for scarce resources, the structure of conventional intermediaries has changed dramatically, incorporating Internet technologies to increase distribution efficiency. In similar fashion, we have begun to see the impact of Internet technologies on the structure of conventional academic journals. However, the change has been relatively slow relative to many other industries. This paper proposes a more efficient approach to the distribution of academic research that builds on the existing strength of academic journals while harnessing the greater efficiency of low-cost internet distribution. It evaluates the effect of the proposed approach by modelling the distributional problems facing academic research, showing how increased distribution efficiency, and the proposed approach in particular, would increase overall research efficiency.

INTRODUCTION

One of the most critical missions of professional academic associations is to promote the efficient development, transfer and utilization of scientific knowledge. This is certainly true of ABSEL. While the Association also addresses the need for personal relationships and networking of like-minded academics and practitioners, these support its larger research mission of promoting and disseminating scientific knowledge.

The problem of knowledge dissemination parallels the larger marketing problem of product or service distribution. Indeed, academic journals, along with other methods through which scientific knowledge is disseminated throughout the scientific community, can be seen as forms of retailing. They collect research from an enormous number of research suppliers throughout the world, evaluate it for quality, and sort it into meaningful assortments that address the needs of different segments of the research community.

Originally, academic journals performed this sorting and assorting themselves. More recently, large electronic data bases such as Ebsco, Jstor and Proquest work with libraries to enable scholars to conduct topical searches and access articles from thousands of different journals without having to deal directly with the journals themselves. Even more recently, Google Scholar has provided still broader access without working through libraries, enabling scholars to conduct similar searches, identifying related articles, and evaluating citation patterns without having to work through libraries.

Even so, the cost of the new distribution systems is enormous. Libraries pay millions of dollars for access to the journals their clients need through the various data bases. Google scholar broadens access by allowing researchers to buy individual articles rather than having to subscribe to full data bases, but the cost of full access to the articles is still very high. Nor does this cost count the time spent by researchers accessing articles, even if they were free. Suppose for instance, that a researcher identifies a number of relevant articles using Google Scholar. Rather than buying them, s/he notes the references and searches for them in the university library's data base. If the journals are not available, s/he contacts the librarians, and they order them through an inter-library network. According to one study, researchers reported spending an average of 26 minutes finding that last article they read when browsing print journals and 40 minutes when browsing electronic data bases (Tenopir, King, Edwards & Wu 2009). In another study, 87.6% of researchers across a broad range of disciplines reported citing 10 or more references per article, with almost half reporting 30 or more. The reported reading an average of 24.3 additional articles for each one cited (Tenopir, Mays & Wu 2011)! Björk, Roos & Lauri (2009) estimate that there were 1,346,000 articles published in 2007 in 23,750 scientific, peer-reviewed journals. While electronic searches take more time than print searches, they also cover a much broader range of potential journals and potentially relevant articles. This increases the quality of the search, but it also multiplies the effect of any fees or time expenditures in retrieving articles.

From the perspective of an organization such as ABSEL whose mission is to promote the development and dissemination of knowledge in its field, the emergence of methods for making relevant research available to large numbers of researchers is exciting. However, the cost is still a major concern. It places a virtual tax on research access, thus dampening the organization's ability to fulfill its mission. The advent of the Bernie Keys Library (BKL), and the larger "classicos initiative" it represents (Cannon & Smth 2004) illustrate the efforts ABSEL has made to address this problem.

The purpose of this paper will be to propose and evaluate an alternative approach to knowledge dissemination that is designed to increase efficiency, and through the savings it evokes, the effectiveness of knowledge creation and dissemination. We will model the approach to determine the conditions under which it would yield an increase in research efficiency.

BACKGROUND

Let us begin by defining the key terms of *effectiveness* and *efficiency.* By *effectiveness* we simply mean the ability to accomplish our objective, which, in this case, is to encourage the development of useful research in our discipline and put it in the hands of the people who need it. *Efficiency* is the amount of useful research produced per unit of resource expended, or in the case of dissemination of knowledge, the amount of useful knowledge disseminated per unit of resource expended. Given the scarcity of resources, efficiency plays an especially important role in the progress of science.

The desire for increased efficiency in disseminating scientific knowledge provides the driving force behind the open-access (OA) movement in scientific publishing. The movement seeks to remove copyright restrictions on scientific publishing, making published research more freely available to scholars all over the world. Early contributrions include Project Muse, initiated by Johns Hopkins University in 1993, now providing full-text open access to a collection of 224,265 articles in 492 journal representing 135 publishers at the time of our writing (Project Muse 2011). In mid-1993, the University of Illinois at Chicago launched an initiative that would not only provide access to journals from other publishers, but would actively search for and disseminate usful content (John 1996).

The current state of the OA movement is perhaps best personified by the Public Library of Science (PLoS), an organization whose specific mission is to

1. Eliminate unnecessary barriers to immediate availability, access, and use of research;
2. Pursue a publishing strategy that drives openness, quality, and integrity;
3. Develop innovative approaches to the assessment, organization, and reuse of ideas and data (PLoS 2011).

In order to address this mission, the PLoS publishes a number of on-line journals. This movement has been supplemented by a number of other open-access or partially open-access publications. For instance, the Directory of Open Access Journals website lists 7,220 journals (DOAJ 2011). The PLoS estimates that open access accounts for roughly 10% of the scientific literature.

In its own way, ABSEL has been a pioneer in the open -access philosophy. The organization pioneered a liberal copyright agreement on its papers. One of the major motivations behind the development of the BKL was to make relevant research papers readily available in an effort to increase the market efficiency of intellectual capital in our field.

Notwithstanding these efforts, ABSEL's primary focus has been to promote research to increase the effectiveness of knowledge development, transfer, and utilization, not the efficiency of its dissemination per se. This is reflected in various reviews of topics addressed at ABSEL conferences, such as Howard and Strang's (2001, 2003) content analysis of the Bernie Keys Library. They cited earlier reviews that identified such themes as the "rigor of research designs and the degree to which they address various educational objectives" (Butler 1985) and "assessment of award-winning procedures and protocols" (Markulis 1991). Their own reviews identified a host of content areas that have been of particular interest to ABSEL researchers. These address our ability to accomplish our objectives, but not how to do it with a lower expenditure of resources.

Looking specifically at research regarding knowledge dissemination, we see the same pattern. Here, studies have addressed such topics as patterns of authorship and citations (Burns & Banasiewicz 1994; Bragge, Thavikulwat, & Töyli 2010) and journals cited in studies in the literature on simulation and gaming (Cannon & Smith 2004; Bragge, Thavikulwat, and Töyli 2010). Again, such studies address effectiveness (insights into how knowledge diffuses), but they offer little insight into the efficiency (cost/benefit) of the process.

An exception is a paper by Cannon & Smith (2010) that was focused specifically on the problem of knowledge dissemination efficiency. The authors begin by characterizing the philosophy and institutional structures supporting academic research in the United States, and increasingly, in the world in general. They argue that efficiency constraints go beyond the obvious barriers created by copyright restrictions and the cost of accessing published research. Their thesis is that the efforts to ensure quality, including those of the open-access movement, are in many cases self-defeating in that they inhibit the efficient dissemination of useful information. That is, the process of editorial review and acceptance or rejection of manuscripts excludes important knowledge from publication. First, the review process is imperfect, resulting in the rejection of many good manuscripts. Second, manuscripts that are fatally flawed can still contain useful knowledge. And third, researchers whose work is unfairly rejected often get discouraged and discontinue their efforts to publish the rejected knowledge.

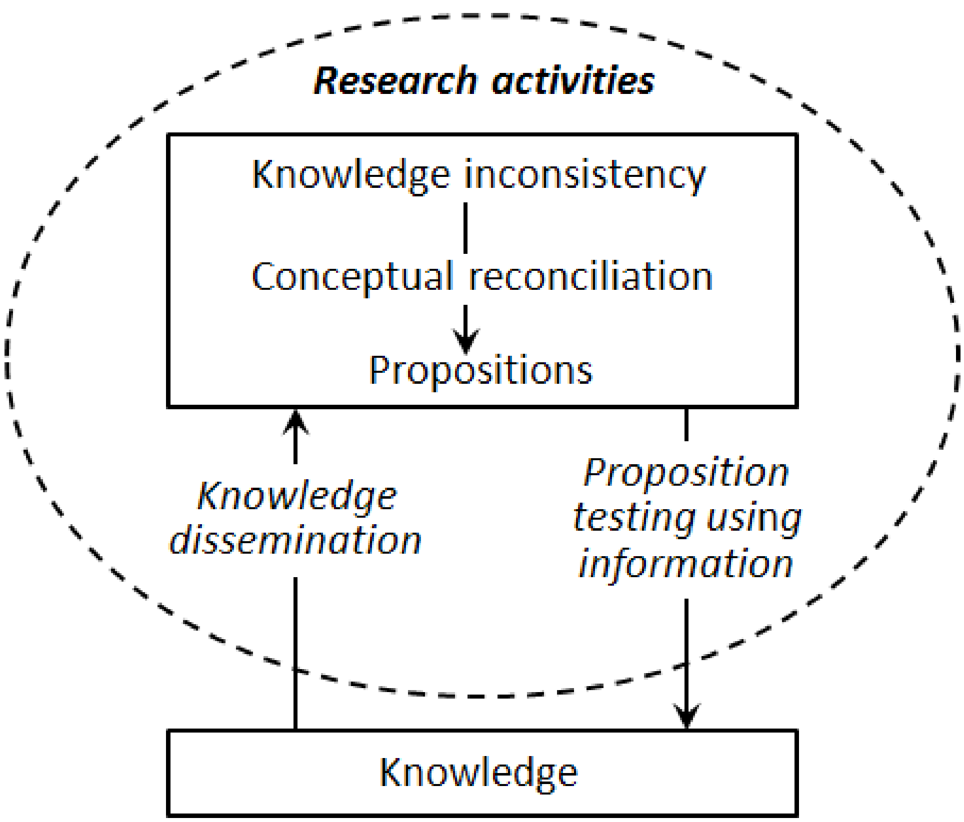
MODELING KNOWLEDGE FORMULATION AND DISSEMINATION EFFICIENCY

In order to develop a more rigorous understanding of the phenomena Cannon and Smith (2010) describe, we propose to construct a more rigorous model of the knowledge formulation and dissemination process. This should enable us to see, and ultimately, to mathematically describe sources of inefficiency. This, in turn, should enable us to determine their relative importance, and, addressing the specific objective of this paper, to evaluate an alternative approach to knowledge dissemination.

The basic tension in the formulation and dissemination of disciplinary knowledge revolves around the two issues discussed in the introduction: (1) How do we maximize the quality and quantity of knowledge created per unit of resources expended on research (research formulation efficiency); and (2) how do we disseminate this knowledge to the people who need it at the least possible cost (knowledge dissemination efficiency). Clearly, these issues are inextricably related in that knowledge dissemination efficiency influences the quantity of resources required for knowledge creation. Likewise, mechanisms used to ensure quality of knowledge creation influences the cost of disseminating knowledge.

Addressing the first issue, an obvious problem in research formulation is the fact that quality and quantity are often inversely related. If we are seeking to optimize the two within a given resource constraint, we need to reduce them to a single measure. We can do this by substituting "relevance" for "quality." We do this by defining knowledge as relevant information, where relevance is determined when information reduces uncertainty regarding the truth of a proposition that is related to a phenomenon under study. Defined in this way, knowledge

Exhibit 1.
The Research Process



is a piece of information that is either relevant or it is not. The quantity of knowledge is simply the sum of relevant pieces of information.

A good way to conceptualize this is through the steps in the scientific method: (1) observation, (2) theorizing, (3) prediction (hypothesis development), (4) testing, and (5) revision in light of the testing results. A *proposition* is a theorized relationship among knowledge components related to some phenomenon encounted in the environment. The relevant *information* is comprised of observations that enable researchers to evaluate (test) the validity of the proposition.

Exhibit 1 portrays our conceptualization. Researchers begin by reviewing available knowledge relating to a particular phenomenon under study. For instance, suppose a researcher was interested in explaining movement in the stock market. S/he would be confronted with a number of seemingly unrelated, and therefore unreconciled, observations. Prices go up, and they go down; some stock values go up, while others go down; and so forth. A conceptual reconciliation might take the form of a proposition saying that prices reflect investor expectations regarding future earnings. Given that these expectations change over time and across companies, the proposition would explain the relationship among the various observations. But is the proposition correct? The proposition may be tested by identifying correlations between relevant information including stock prices and forecasts of earnings for the companies. The combined proposition and supporting information would constitute a contribution to our knowledge.

Now, consider the impact knowledge has on the formulation of new propositions. For instance, suppose that another researcher found that the variation in stock prices tended to be greater than we would expect based solely on the changes in company earnings. This might lead to the proposition that investors tend to be overly optimistic and pessimistic, depending on the direction the market is going. The new proposition can be tested using investor perception information, creating more knowledge. And so the process continues.

The critical elements of this analysis are twofold: First, the formulation of knowledge plays an important role in stimulating the formulation of more knowledge. Second, no single researcher is doing all the research, so the dissemination of knowledge from one researcher to another becomes a key element in process of knowledge formulation. Again, note that the purpose of this discussion is to develop a rigorous conceptualization of how knowledge grows. Ultimately, this can be reduced to mathematical form so that we can develop testable propositions regarding the effectiveness and efficiency of new methods of knowledge dissemination. We will propose and evaluate one such model in this paper.

The concepts portrayed in Exhibit 1 lend themselves to relatively simple mathematical notation. If K represents the aggregate quantity of relevant knowledge, then we can represent it in the form of Equation (1):

$$K = \sum_{i}^{m}\sum_{j}^{n} K_{i,j} = \sum_{i}^{m}\sum_{j}^{n} I_i \cdot P_{i,j} \qquad (1)$$

I_i represents a specific unit of information and $P_{i,j}$ a phenomenon regarding which the information reduces uncertainty. $K_{i,j}$ represents a potential unit of relevant knowledge, the product of I_i and $P_{i,j}$. Thus, if a unit of information exists (I_i=1) *and* there is a corresponding phenomenon for which it reduces uncertainty ($P_{i,j}$ =1), we have a unit of *relevant* knowledge ($K_{i,j}$=1). Conversely, if either the required information *or* the phenomenon for which it is needed to reduce uncertainty is missing (I_i=0 or $P_{i,j}$ =0) there is no unit of relevant knowledge ($K_{i,j}$=0). K is the sum of all units of relevant knowledge.

The efficiency of knowledge formulation (E_K) is captured in Equation (2):

$$E_K = \frac{K}{C_K} \qquad (2)$$

where C_K is the cost of resources required to produce K units of relevant knowledge.

D represents knowledge dissemination, or the total amount of relevant knowledge delivered to researchers who need it. This is represented by Equation (3):

$$D = \sum_{i}^{m}\sum_{j}^{n} D_{i,j} = \sum_{i}^{m}\sum_{j}^{n}\sum_{k}^{z} K_{i,j} \cdot \frac{R_{i,j,k}}{z_j \mid z_j \neq 0} \qquad (3)$$

$R_{i,j,k}$ represents a researcher k who is studying phenomena related to proposition j and is therefore in need of knowledge $K_{i,j}$. If $R_{i,j,k}$ =1, $K_{i,j}\times R_{i,j,k}$ takes on the value of one when the researcher has retrieved $K_{i,j}$ from the body of existing research. According to the equation, if $K_{i,j}$=0, we have no relevant unit of knowledge to disseminate, so $D_{i,j}$=0. If $K_{i,j}$=1, but $R_{i,j,k}$=0, the unit of relevant information has not been effectively disseminated or there is no researcher interested in relevant knowledge $K_{i,j}$, so there is still no contribution to D. If both $K_{i,j}$=1 and $R_{i,j,k}$=1, we have an incident of knowledge dissemination.

In practice, we know that more than one researcher is likely to be interested in a given piece of relevant knowledge. Therefore, to be effectively disseminated, the knowledge would have to reach all of the interested researchers. We represent this in Equation (3) by multiplying $K_{i,j}$ by the proportion of interested researchers to which the knowledge has been disseminated. The proportion is found by dividing $R_{i,j,k}$ by z_j, the total number of researchers interested in phenomenon j. To illustrate, if 20 researchers are interested in phenomenon j, but only 10 these are exposed to relevant knowledge $K_{i,j}$, the knowledge would be disseminated to half the researchers who need it ($D_{i,j}$=0.5). Of course, the proportion is only meaningful if there is at least one researcher interested in studying the phenomenon (i.e. z_j must not equal to zero).

The net effect of Equation (3) is to give D a value from zero to K.

If C_D represents the total cost of knowledge dissemination, Equation (4) would represent E_D, the efficiency of knowledge dissemination:

$$E_D = \frac{D}{C_D} \tag{4}$$

E_D, however, is an inadequate measure. We could increase E_D by simply ignoring difficult-to-reach researchers. D would be low, but so would C_D, so efficiency would be high. The poblem, of course, is that knowledge dissemination would be ineffective, leaving large numbers of researchers without the knowledge they need to pursue their research. A more meaningful measure market effectiveness would consider knowledge dissemination effectiveness *relative* to the total value of K. We can formulate a measure, M, where

$$M = \frac{D}{K} \tag{5}$$

M =1 would characterize a market where every relevant piece of knowledge is disseminated to the people who need it. M=0 would characterize a totally ineffective market in which no knowledge ends up in the hands of the people who need it. The resulting measure of efficiency would be the proportion of knowledge distributed per unit of resources expended on knowledge dissemination, or

$$E_M = \frac{M}{C_D} \tag{6}$$

Ideally, market pressures would optimize C_D such that the marginal value of increased knowledge dissemination equals its marginal cost. And in fact, the scientific journal publication market is dominated by private enterprises whose efficiency should be controlled by market forces. Nevertheless, Cannon and Smith (2010) argue that the market suffers from major imperfections, many of which might be addressed by a new approach. However, an additional component of our knowledge-dissemination system is comprised of the distribution channels for scientific journals. These are largely funded by library budgets, which are, in turn, supported in large part by public funds, which are primarily determined by political rather than market forces. We address this by assuming that a proportion of C_D is fixed.

THE RELATIONSHIP BETWEEN KNOWLEDGE FORMULATION AND DISSEMINATION

Looking beyond this study, our overriding interest is in knowledge formulation, not dissemination. Knowledge dissemination is only relevant insofar as it stimulates a growth in knowledge. Figure 1 provides the link. The process is based on the premise that knowledge builds on itself over time. According to Sorenson & Fleming (2004), this premise is supported by three streams of research. First, macroeconomic studies have associated increases in scientific employment (Sveikauskas 1981) and increased expenditures on research and development (Mansfield 1972; Adams 1990) with growth in GDP. Second, looking at firm-level data, studies have shown that companies that nurture professional research activities (attending conferences, publishing papers, etc.) tend to outperform firms that do not (Henderson & Cockburn 1994). The third stream looks at actual knowledge production, using citation analysis to track the impact of published knowledge on future inventions (Jaffe & Trajtenberg 1996; Mowery & Ziedonis 2002; Sorenson & Fleming 2004). In these studies, "inventions" are operationalized as patents. Following this model, one could address the generation of knowledge by simply substituting scientific publications for patents.

If we accept that knowledge builds on itself, and that knowledge increases continually, we would expect an exponential growth of K over time. The actual rate of growth would depend on dissemination effectiveness (M) and growth in the number of researcher-projects being conducted (R), where R is derived from Equation 3, as shown in Equation 7:

$$R = \sum_{i}^{m} \sum_{j}^{n} \sum_{k}^{z} R_{i,j,k} \tag{7}$$

Recall that $R_{i,j,k}$ represents a researcher k who is studying phenomena related to proposition j and is therefore in need of knowledge $K_{i,j}$. Given this notation, Equation 8 represents the growth of knowledge over time:

$$\Delta K_t = f(M, R) \tag{8}$$

where Δ represents change from time t-1 to time t.

The equation suggests that we might increase knowledge growth in two ways: (1) by increasing effectiveness of knowledge dissemination (M), and (2) by increasing the total number of research projects being conducted (R). In order to understand the trade-offs between M and R, we need to consider the effects of resource constraints on Equation 8. If resources were no object, we would find ways to distribute all knowledge, so D=K and M=1. The effect would be to increase knowledge growth (ΔK_t). Researchers would be able to take on more projects with greater success in creating knowledge. Assuming that researchers are rewarded for the knowledge they create, salaries would rise and workers would flock to become researchers. Again, assuming no resource constraint, research institutions would continue to hire researchers. As we impose constraints, research institutions would hire researchers until diminishing returns in their productivity off-set the effect of perfect knowledge dissemination, reaching a point of equilibrium. Similarly, diminishing returns on knowledge dissemination would cause *M* to settle at some level lower than 1 as we impose resource constraints on it. Let us begin by establishing a budget constraint, *B*, including separate budgets for knowledge formulation (C_K) and dissemination (C_D), as shown in Equation 9:

$$B \geq C_K + C_D \tag{9}$$

Equation 10 and Equation 11 represent the effects the budget constraints have on *D* and *R*, respectively:

$$D = g(C_D) \tag{10}$$

$$R = h(C_K|D) \tag{11}$$

Functions *g* and *h* are expressions of knowledge-dissemination and knowledge-formulation efficiency, respectively. Function *g* represents the response of knowledge dissemination to allocation of resources (C_D). Function *h* represents the corresponding response in research effort to allocation of resources (C_K). Specifically, researchers will engage in a research project when the value derived from creating knowledge is greater than or equal to the cost of creating it (C_K), given some level of knowledge dissemination. Increasing dissemination makes the knowledge needed to conduct research relatively more available, thus decreasing the cost of research and enabling researchers to take on more research projects. In other words, the number of research projects (*R*) is an indirect function of knowledge dissemination *D*), as shown in Equation (11).

While we don't know the exact functions represented by *g* and *h*, we have some evidence that bears on their relative shapes. We are discussing primarily academic researchers, where we can assume that the number of researchers changes very slowly. This is inherent in the tenure system and the self-perpetuating nature of academic institutions. Presumably, such changes as do occur would be strongly influenced by levels of funding (C_K). There is evidence, however, that changes in funding levels have relatively little impact on the rate of knowledge growth (Cohn 1986). If this is true, then the primary driver of *R* would be *D*. Again, this suggests that knowledge dissemination is the key element of Equation (11), and by extension, Equation (8). Sorenson & Fleming (2004) offer even more direct evidence regarding the importance of knowledge dissemination. They demonstrated that patents referencing articles from peer-reviewed scientific journals resulted in a greater growth in knowledge (as reflected in follow-on patents) than those that referenced commercial (non-peer-reviewed) publications or that referenced no publications at all. In the same vein, Boyce, King, Montgomery, and Tenopir (2004) and Tenopir, King, Spencer, & Wu (2009) found that the most published researchers and those that have won awards tend to read more scholarly journals than their less productive colleagues.

ADDRESSING THE PROBLEM OF RESEARCH QUALITY AND RELEVANCE: THREE ALTERNATIVE APPROACHES

Now let us return to the question of research relevance. Recall that we used the concept of relevance to combine our measures of quantity and quality of research. We defined knowledge as relevant information, where relevance was determined when information reduced uncertainty regarding the truth of a proposition that is related to a phenomenon under study. This enabled us to determine the quantity of knowledge by simply adding up the units of relevant information.

This formulation assumed that the unit of information (I_i) that formed the basis for knowledge was a dichotomous variable. That is, it either existed or it did not; it represented a proposition that was either true or false. But what if we are not sure? The issue of *quality* revolves around the probability that a proposed unit of relevant information is true. The higher the probability, the higher the quality of the research that proposed the information, expressed as values of I_i btween 0 and 1.

We can think of relevance in the same way. Rather than treating $P_{i,j}$ as a dichotomous variable, where a piece of information will either reduce uncertainty or not reduce uncertainty regarding the proposition, we can view it as the probability that the information will reduce uncertainty. $K_{i,j}$ (as developed in equation 1) would represent the joint probability of truth and relevance.

So, how do we evaluate this joint probability? The most common way is to evaluate the quality of the journal in which the research is published. Journal quality functions as a proxy for the quality and relevance of the knowledge published in the journal. Of course, journal quality is a measure of perception. In the end, this perception -- a journal's reputation – tends to depend on the quality and relevance of the research it publishes, managed through its editorial and review processes.

In our earlier discussion of Equation (4), we introduced the cost of knowledge dissemination (C_D). This can be further broken down into the cost of publication (C_P – cost of making the research accessible to potential users) and the cost of review (C_R – cost of determining the quality and relevance of the research). The value of C_D is represented in Equation (12).

$$C_D = C_P + C_R \tag{12}$$

While the actual cost of publishing (C_P) no doubt varies dramatically by discipline, journal, and journal type, Odlyzko (1997) estimates the total "systems" cost for a single journal article as follows, based on a sample of mathematics and computer science journals:

- revenue to the publisher: $4,000 (primarily library subscriptions)
- library costs other than subscriptions (i.e. "overhead"): $8,000
- authors' costs of preparing a paper: $20,000
- editorial and refereeing costs: $4,000

The most obvious avoidable cost is revenue to the publisher. The $4,000 represents the total revenue received from a journal, divided by the number of published articles. While this revenue pays for the services rendered by the publisher, reviewing a number of different studies and publishing scenarios, Odlyzko concludes that virtually all estimates for an electronic journal would fall between $300 and $1,000 per article.

The other costs are harder to analyze. However, library overhead would certainly be reduced if more information were available through independent web posting. Even more dramatic would be the savings in the cost to authors preparing papers. Consider the process through which researchers conduct a literature review. One of the most efficient ways to identify relevant articles is through Google Scholar. Once a scholar has identified a potentially relevant article, s/he can access it directly with a single click. However, the royalties will usually range between $20 and $40 per article, and the time required for effective access multiplies several times due to the simple logistics of making online payment. If the scholar has access to a subsidized university library, the monetary cost goes away, but the time required to access the articles increases even more.

We found no studies indicating the actual costs of information retrieval using Internet-based search engines and data base access for conventional versus open-access electronic journals. However, by making some conservative assumptions, we can develop a "ball-park" estimate of the difference. Recall that, in one study, almost half of the researchers studied cited 30 or more articles, and that they retrieved an average of 24.3 additional articles for each one cited (Tenopir, Mays & Wu 2011). Assuming an article has 30 citations, this requires the retrieval of (30 x 24.3=) approximately 728 articles. Now, assume that a researcher uses Google Scholar to conduct an initial literature review, delegating the actual retrieval of articles to a graduate research assistant (GRA). Ignoring the cost of time taken to print a list of potentially interesting articles, assume that the GRA requires five minutes to retrieve each article, the task requires (728 x 5/60=) approximately 60 hours. Assuming a total cost per hour for the GRA of $30, the cost would be (60 x $30=) $1,800 per journal article more for accessing conventional versus electronic, open-access journal, just for information retrieval.

Of courses, this assumes that the researcher has access to a library through which a data base containing the required journal articles is available. In an increasingly global market, we find an increasing number of scholars engaging in research who do not have access to such libraries. The potential contribution of these researchers is enormous. However, the costs they face to deliver the same quality of research as their better-supported colleagues is enormous.

If we are successful in making research available online in open-access format, changes in the rate at which the cost of knowledge dissemination decreases becomes a function of the cost of determining the quality and relevance of research. This begs the question, how cost effective are review processes in determining quality and relevance? Drawing on Cannon and Smith (2010), we can identify three prototypic approaches to review:

- *Ad hoc review*. *Ad hoc review* is the traditional journal review system. Researchers submit manuscripts to an editor who, in turn, evokes the reviews of experts regarding its quality. Based on these reviews, researchers may be asked to revise their manuscripts, but ultimately, the editor makes a decision to publish or reject it. This is also referred to as an "exclusionary" review approach because the review process excludes articles from publication that are not judged to be of high enough quality and/or relevance to merit publication in the journal conducting the review.
- *Post hoc review*. By contrast, *post hoc* review approaches involve the publication of manuscripts that their authors judge to be significant, leaving the evaluation of quality to some form of review after they are published. Typically, this publication would be on the world-wide web, and the review process would involve some form of intellectual market feedback (such as, but not necessarily limited to citation rates, comments/replies, and/or ratings). This is also referred to as an "evaluative" review approach because the reviews simply evaluate articles without determining their publication.
- *Hybrid review*. Hybrid review approaches, as the name implies, are some combination of the first two. That is, they involve both *ad hoc* and some form of *post hoc* review. To illustrate, in its weakest form, the evaluation of traditional peer-reviewed journals or journal articles by various forms of citation analysis is a hybrid model, where the journal content is determined by an exclusionary process, but the journal/ article quality is partially determined by *post hoc* review. An intermediate form would be illustrated by the common practice of open-access posting of manuscripts on-line, then removing them once they appear in a journal. Finally, the strongest form would be illustrated by simultaneously open-access posting of a paper, or its contents, and the actual manuscript in an exclusionary journal.

The *ad-hoc review* approach minimizes the probability that low-quality research is disseminated. That is, knowledge is withheld from researchers until there is sufficient assurance that the knowledge is of high quality. This is beneficial in that it increases the probability that the research yields relevant knowledge (increasing the value of $K_{i,j}$). However, the *ad-hoc* approach also reduces the total number of potentially relevant units of information (m) available to researchers, thus limiting the total the potential for testing propositions, formulating knowledge, and stimulating the development of new propositions (as illustrated in Exhibit 1).

In contrast, the *post-hoc review* approach maximizes the total number of potentially relevant units of information available (m) for dissemination. However, the probability that the research is of sufficient quality and relevant may not be ascertained until after the knowledge is disseminated. This review process essentially imposes the cost of review on the body of researchers interested in a given phenomenon.

While this cost impedes knowledge formulation in the beginning (as compared to knowledge whose quality and relevance is already established at the time of publication through *ad hoc* review process), this "beginning" comes before that of conventional journal articles, where the review process can delay publication many months, or even years, beyond the time of submission. By contrast, *post hoc* review can achieve high levels of effectiveness in a very short period of time.

One of the best examples of *post hoc* review is the "wiki" movement in information dissemination (Tapscott & Williams 2006) where self-selecting consumers of information evaluate the relevance of information. One study of change histories in Wikipedia articles found that "malignant" edits were corrected in a matter of minutes after posting. "Malignant" posts represented a kind of intellectual vandalism – a deliberate attempt to distort the truth. We might infer that these posts were very important to members of the Wikipedia audience, thus motivating high vigilance and a rapid response. However, the speed with which they were corrected illustrates the potential for the user-based *post hoc* review approach.

For research articles, presumably the self-selection of reviewers is based on having encountered an article as part of a literature review relating to a phenomenon being researched. Therefore, much of the cost of the reviewing the article for quality and relevance is already accounted for in the cost of preparing a specific manuscript. This is much more efficient than mobilizing members of a journal's editorial board to conduct a manuscript review in the traditional journal reviewing system. In other words, we would achieve economies by sharing costs between C_P and C_R.

Using law reviews as an example, Hibbits (1996) argues for such a process, suggesting that even electronic journals are inefficient in an era of web-based self-publishing. While he considers a number of arguments for and against the self-publishing/post-hoc review approach, they can be reduced to the three general issues we have identified: (1) quality, I_i, (2) relevance, $P_{i,j}$, and (3) effectiveness of knowledge dissemination (M), all of which have been central to our discussion so far. Again, the "poster-child" for post-hoc review is Wikipedia. Addressing issue (1), Chesney (2006) provides evidence that Wikipedia's post-hoc review process produces relatively high levels of accuracy in postings. An often-cited article in *Nature* (Giles 2005) reports that the accuracy of Wikipedia is very comparable with that of Encyclopaedia Britannica. Addressing issue (2), Banerjee, Ramanathan, & Gupta (2007) show that the classification algorithms that link articles in Wikipedia and Google do an excellent job of identifying clusters of related (relevant) information.

Addressing issue (3), we have already noted that *ad hoc* review inherently reduces the number of units of information (m), and hence, the amount of knowledge available to researchers. As noted earlier, Cannon and Smith (2010) identify three key elements of this problem: First, the review process is imperfect, resulting in the rejection of many good manuscripts. Second, manuscripts that are fatally flawed can still contain useful knowledge. And third, researchers whose work is unfairly rejected often get discouraged and discontinue their efforts to publish the rejected knowledge.

The arguments in favor of a *hybrid* approach are threefold: First, maintaining the existing journal structure capitalizes on the existing culture and reward systems of the scholars who produce research, as well as its functional contributions. Nowithstanding the advent of electronic search engines, browsing journals for relevant articles is still a very popular method of information retrieval (Boyce, King, Montogery & Tenopir 2004; Tenopir & King 2008), suggesting that journals do a relatively good job of screening for quality and relevance. Second, aside from the evaluative process that goes into the acceptance of articles for journal publication, the ad hoc review process also provides important feedback for improving research. Third, adding a *post hoc* evaluative process allows interested researchers to further determine research quality, signal interest in establishing relationships, and to promote additional propositions.

Our purpose in this paper will be to present and evaluate a semi-strong-form *hybrid* model where the knowledge formulation market supports open-access peer reviews of articles (literature reviews, topical surveys, and/ or editorials) and actual articles published in exclusionary journals. We will begin by presenting the model and the basic logic behind it. We will then explain the expected impact it will have on knowledge dissemination effectiveness, knowledge dissemination efficiency, and ultimately the rate of knowledge growth. We will conclude by discussing the model in light of the practical realities of the modern academic research establishment.

A SEMI-STRONG-FORM HYBRID APPROACH TO RESEARCH EVALUATION

Let us return to the concept of journals as intermediaries in the market for scientific knowledge, as discussed in the introduction to this paper. Their economic justification rests in the functions they perform, gathering articles from researchers all over the world, screening them for quality, facilitating improvements, and publishing them in collections that other researchers will find accessible and relevant to their work.

We have noted the role that online-search engines play in the knowledge-dissemination process. There were approximately 1,350,000 articles published in 23,750 scientific, peer-reviewed journals during 2007 (Björk, Roos, & Lauri 2009), a number that continues to rise each year. With it, the manner in which scholars conduct research is also changing. Rather than browsing journals, researchers are increasingly using search engines to find articles that more precisely address the phenomena they are studying (Boyce, King, Montgomery, & Tenopir 2004; Tenopir & King 2008).Returning to Exhibit 1, we would expect this to have two effects on research productivity. First, increasing the number of relevant research articles to which researchers are exposed would confront them with increasing levels of unreconciled knowledge. This, in turn, would stimulate them to formulate an increased number of research propositions. Second, they would be more likely to find specific studies that provide the information needed to address the propositions they have formulated. This would obviate the need for many relatively more expensive empirical research projects.

The changing research environment highlights the relative importance of knowledge dissemination efficiency. Using the notation we have developed in our previous equations, the amount of research done (R) per dollar spent (C_K) depends on the level of knowledge dissemination (D), as suggestion in Equation (11). The level of D, in turn, depends on the budget allocation to knowledge dissemination (C_D). As the price per research article retrieved decreases, both in dollars and in the time it takes to retrieve articles once they are located, the entire response curve represented by function g shifts, lowering the marginal cost of D. This, in turn, shifts the response curve represented by function f, lowering the marginal cost of knowledge (K). The effect is a leveraged acceleration of knowledge growth, which, as we have noted, is not only a national priority, but also ABSEL's primary objective as an organization.

Our "modest proposal" seeks to capitalize on the efficiency of Internet communication and the OA movement while retaining the functions of conventional journal publication. It introduces a "pseudo-journal" where researchers provide synopses and critical reviews of journal articles. This is consistent with the proactive approach taken at the University of Illinois at Chicago, where the university library decided to "[scout] potential sources of information in an aggressive campaign to find organizations and individuals with important, useful information that the Library could develop and deliver over the Internet" (John 1996).

Using a retail analogy, the content of this journal would be similar to that of a consumer report, including a description of an article, comparisons with similar articles, and an assessment of quality. Our model proposes that the "pseudo-journal" be publicly available for interested researchers to read and contribute.

First, the "pseudo-journal" articles have a direct effect on knowledge quality. Articles that assess the quality of journal articles increase the certainty that any given article has reliably matched truth (I_i=1) with a relevant proposition (Pi,j=1). The increased certainty increases the expected value of aggregate relevant knowledge (K, as shown in Equation 1).

Second, "pseudo-journal" articles will have an indirect effect on knowledge quality. The *post hoc* quality assessment process would influence the ad hoc review process component of the *hybrid* model. As discussed earlier, private journals engage in rigorous review processes to establish perceived journal quality, which in turn, attracts researchers because the expected value of published relevant knowledge (K) is higher than in other journals. A *post hoc* assessment of published article quality may increase or decrease a journal's perceived quality. Journals' editorial boards are likely to increase the rigor of their review processes knowing that premier journals are likely to receive significant attention in "pseudo-journal" articles. More rigorous review processes increase the certainty that knowledge has accurately matched truth with a relevant proposition ($\uparrow K_{i,j}$), increasing the expected value of aggregate relevant knowledge (K).

Third, referencing Figure 1, the "pseudo-journal" influences conceptual reconciliation element of the research process. Researchers can use the "pseudo-journal" to assemble existing knowledge to inform and formulate propositions. To illustrate, let us return to the phenomenon of movement in the stock market. Suppose one or more researchers publish an article in the "pseudo-journal" that describes various research articles relating to expected future earnings. Specifically, the article summarizes published research that suggests stock prices reflect investor expectations about future earnings and that current earnings provide information about expected future earnings. These two pieces of knowledge beg the proposition that stock price changes are associated with changes in current earnings, which in turn, begs other propositions that stock price changes are associated with operational factors that influence current earnings. By publicly assembling categorical knowledge, the "pseudo-journal" increases knowledge dissemination effectiveness (M) for researchers ($R_{i,j,k}$) interested in that particular phenomenon. Greater dissemination effectiveness increases the number of research projects (*R*) that investigating researchers can manage (equation 11), which in turn increases the aggregate knowledge growth rate (equation 8).

Finally, the "pseudo-journal" influences researchers' knowledge search patterns. Existing dissemination mechanisms such as Google Scholar produce brief synopses of potentially relevant journal articles. In addition to synopses, "pseudo-journal" articles are intended to provide comparisons between potentially relevant articles. The comparisons inform researchers' search patterns, focusing their use of resources on articles that are most relevant to the phenomenon they are interested in. In our mathematical model, focused searches are represented by reducing the quantity of resources (C_D) required to dissemination knowledge (*D*), increasing knowledge dissemination effectiveness (*M*) and aggregate knowledge growth rate (equation 8).

SUMMARY AND CONCLUSIONS

The advent of the Internet and associated information-dissemination technologies has stimulated an enormous amount of soul-searching in the research establishment. Is there some way to capture the potential distribution capabilities of the Internet to increase the efficiency of scientific knowledge dissemination? ABSEL has been an enthusiastic participant in this search for efficiency. The development of the BKL is one manifestation. However, manner in which the BKL archives should be administered, along with other potential knowledge-dissemination initiatives, continues to be a topic of discussion that is central to our organization.

Of particular interest is a paper presented at the 2010 conference by Cannon and Smith that evaluates alternative approaches to the publication problem. This paper builds on their work, suggesting a specific instantiation of what they call a "hybrid" review approach, combining both the conventional ad hoc review process of conventional academic conferences and journals and the emerging post hoc review process typified by Wikipedia and related vehicles for self-publishing research results.

In order to develop a rigorous framework from which to evaluate alternatives, we have sought to model the key elements of the knowledge distribution system. First, we formulate a rigorous definition of the knowledge formulation process, showing how knowledge builds on itself to create propositions, and how it reduces uncertainty with respect to these propositions, creating new knowledge. Second, we show how knowledge distribution impacts on the knowledge creation process, articulating some major barriers that impede efficient knowledge formulation. Finally, having articulated the economic functions of our existing system of academic research publication (focusing on the role of peer-reviewed journals), we suggest a complementary approach to publishing in which researchers (or their surrogates) publish detailed synopses of their work and the works of others to make them more accessible at a lower cost. We show how the proposed system interacts with conventional journals to capitalize on their strengths (publishing high-quality, relevant research) while compensating for their weaknesses (increasing the cost of knowledge dissemination and diminishing the overall quantity of knowledge published).

We argue that both the conventional and our suggested "self-publishing" approach are subject to market pressures for efficiency and effectiveness. We believe that, over time, the new approach will gain prominence. However, by creating a framework for peaceful coexistence, and even synergy, the combined approach capitalizes on the strengths of both *ad hoc* and *post hoc* review processes, allowing the market system to move at its own pace and, ultimately, in the direction researchers find most beneficial to their work.

Clearly, an important agenda is to offer our economic model as a framework for future research in the area of system-wide ("macro") analysis of knowledge formulation and dissemination effectiveness and efficiency. For example, our economic model can be used to frame the impacts of researcher incentives (e.g. promotion and tenure processes) on knowledge formulation and dissemination effectiveness and efficiency. Few subjects are more central to the work we do as researchers, or more specifically, as an association of researchers dedicated to the creation and dissemination of knowledge related to business simulations and experiential learning. The model not only leads to additional propositions, but casts them in a framework that lends itself to empirical testing.

REFERENCES

Adams, J. D. (1990). Fundamental Stocks of Knowledge and Productivity Growth. *Journal of Political Economy* 98(4), 673-702.

Bannerjee, S., Ramanathan, K., & Gupta, A. (2007). Clustering short texts using Wikipedia. *Proceedings of the 30th annual international ACM SIGIR Conference on Research and Development in Information Retrieval*.787-788.

Boyce, P, King, D. W., Montgomery, C. & Tenopir, C. (2004). How Electronic Journals Are Changing Patterns of Use. *The Serials Librarian*, 46(1-2), 121-141.

Bragge, J., Thavikulwat, P., & Töyli, J. (2010). Profiling 40 Years of Research in Simulation & Gaming. *Simulation & Gaming*, 41(6), 869-897.

Burns, A. C. and Banasiewicz, A.(1994). The Intellectual Structure of ABSEL: A Bibliometric Study of Author Cocitations over Time. *Developments in Business Simulation & Experiential Exercises*, 21, 7-12. Reprinted in the *Bernie Keys Library*.

Björk, B. C., Roos, A., & Lauri, M. (2009). Scientific

journal publishing: Yearly volume and open access availability. *Information Research*, 14(1) [online].
Cannon, H. M. and Smith, J. A. (2004). The ABSEL Classicos Initiative. *Simulation & Gaming*, 35(2), 294-307.
Cannon, H. M. and Smith, J. A. (2010). Strategies for promotiong knowledge formulation and dissemination in business simulation and experiential learning. *Developments in Business Simulation & Experiential Exercises*, 37, 258.266. Reprinted in the *Bernie Keys Library*.
Chesney, T. (2006). An Empirical Examination of Wikipedia's Credibility. *First Monday*. 11(6) [online]
Cohn, S. F. (1986). The Effects of Funding Changes upon the Rate of Knowledge Growth in Algebraic and Differential Topology, 1955-75. *Social Studies of Science*, 16(1), 23-59.
DOAJ (2011). http://www.doaj.org/ [Accessed October 25, 2011]
Giles, Jim (2005). Internet Encyclopaedias Go Head to Head. *Nature* 438, 900-901.
Henderson, R. and Cockburn, I. (1994). Measuring Competence? Exploring Firm Effect in Drug Discovery. *Strategic Management Journal*, 15(), 63-84.
Hibbitts, B. J. (1996). Last Writes? The law Review in the Age of Cyberspace. *First Monday*, 1(3) [online].
Howard, B. & Strang, D. (2001). ABSEL: They Way We Talk!, *Developments in Business Simulation & Gaming*, 28, 111-116. Reprinted in the *Bernie Keys Library*.
Howard, B. & Strang, D. (2003). The Way We Talk! Take II, *Developments in Business Simulation & Gaming*, 30, 133-145. Reprinted in the *Bernie Keys Library*.
John, N. R. (1996). Putting Content onto the Internet. *First Monday*, 1(2). [Online]
Larsen, P. O. & von Ins, M. (2010). The Rate of Growth in Scientific Publication and the Decline in Coverage Provided by Science Citation Index. *Scientometrics*, 84 (3), 575-603.
Leisyte, L., Enders, J. & de Boer, H. (2009). The Balance between Teaching and Research in Dutch and English Universities in the Context of University Governance Reforms. *Higher Education* 58(5), 619-635.
Mansfield, E. (1971). Contribution of R&D to Economic Growth in the United States. *Science*, 175(4021), 477-486.
Odlyzko, A. (1997). The Economics of Electronic Journals. *First Monday*, 2(8). [online].
PLoS (2011). PLoS Expands Mission. http://www.plos.org/plos-expands-mission/#more-568 [Accessed October 25, 2011].
Project Muse (2011). Current Collection. http://muse.JHU.edu [Accessed November 10, 2011].
Sorenson, O. & Fleming, L. (2004). Science and the Diffusion of Knowledge. *Research Policy*, 33(10), 16154.1634.
Sveikauskas, L. (1981). Technological Inputs and Multifactor Productivity Growth. *Review of Economics and Statistics*, 63(2), 275-282.
Tapscott, D. & Williams, A. D. (2006). *Wikinomics: How Mass Collaboration Changes Everything*. Portfolio: New York, NY.
Tenopir, C. & King, D. (2008). Electronic Journals and Changes in Scholarly Aritlce Seeking Reading Patterns. *D-Lib Magazine*, 14(11-12) [online].
Tenopir, C., King, D. Edwards, S. & Wu, L. (2009). Electronic Journals and Changes in Scholarly Article Seeking and Reading Patterns. *Aslib Proceedings* 61 (1), 5-32.
Tenopir, C., King, D., Spencer, J. & Wu, L. (2009). Variations in Article Seeking and Reading Patterns of Academics: What Makes a Difference. *Library & Information Science Research*, 31(3), 139-148.
Tenopir, C., Mays, R. & Wu, L. (2011). Journal article growth and reading patterns. *New Review of Information Networking*. 16(1), 4-22.
Viégas, F. B., Wattenberg, M. & Dave, K. (2004). Studying Cooperation and Conflict b etween Authors with *History Flow* Visualizations. *SIGCHI Conference on Human Factors in Computing Systems*, 24-29April, Vienna, Austria, 575-582.
Vincent-Lancrin, S. (2006). What Is Changing in Academic Research? Trends and Futures Scenarios. *European Journal of Education* 41(2), 169-203.