The Potential for Scholarly Online Books: *Views from the Columbia University Online Books Evaluation Project**

Mary Summerfield, Carol Mandel, and Paul Kantor

The Columbia University Online Books Evaluation Project sought to understand both user reactions to online books in the scholarly world and the cost profiles of print and online books. Scholars appreciated the opportunity to use the online format to locate a book and to browse it. However, they sought a print copy for extended reading. Incremental costs of online books are small for publishers. Libraries' life-cycle costs are lower for online books than for print books.

Introduction

From 1995 to 1999, the Online Books Evaluation Project at Columbia University studied online books as resources in the academic world. This article summarizes key findings on the potential role of online books in the scholarly world and on various marketplace issues. *The Online Books Project, Columbia University: Final Report* and many other reports and papers, including the research protocol, are available at *http://www.columbia.edu/cu/libraries/digital/texts/about.html.*

The project looked at (1) scholars' behavior and reactions to online books, (2) lifecycle costs of traditional print books and online books for publishers and libraries, and (3) marketplace reactions to the concept of online books. All of these were assessed in the context of environmental developments in the relevant areas of academia, scholarly communication, and the computing and Internet worlds. The evaluation employed a wide variety of tools, including server data, online, mailed, and hand-distributed surveys, and individual and group interviews.

*Summerfield and Mandel were both members of the Columbia University Libraries staff during the time of this project. Summerfield was a Project Director and Coordinator of this Project and Mandel was Deputy University Librarian and Director of this Project.

Development and Design of the Online Books Collection

The project was launched in January 1995. Initially we expected to develop custom SGML browsers, as other online publishing projects were doing. However, at launch we judged the Web to be the best delivery system, for two main reasons. First, scholars would have access to the Web from university locations and many, if not most, would be able to access it from home. Second, the Web would maximize the value to scholars, as the greatest added value from online books would come with truly digital books.[1] Perhaps only such enhanced online books would offer sufficient advantages over the traditional print format that scholars would choose to substitute them for the print format in some or all of their modes of use and for some or all classes of books.[2]

As of June 1999, the online book collection contained 168 texts, including six reference works and 54 classical texts in social thought. Four publishers provided modern books to the collection—Columbia University Press, Oxford University Press, Garland Press, Simon and Schuster Higher Education. These contemporary works were in six subject areas: biography, literary criticism, earth and environmental science, philosophy, political science/international affairs, and social work. A few of these books were designed as college textbooks; others were scholarly monographs or collections of essays. Each book was in the Libraries' collection in print form as well as in one or more online formats.

At the outset, the project editorial team decided to mount books with each major element as a separate file, e.g., a table of contents-title page, the introduction, each chapter, the bibliography, the index. The user could move to a chapter by clicking on its title in the table of contents. Footnote numbers were linked to the footnote contents; a *back* link returned the reader to the text. To move among chapters the scholar returned to the table of contents and clicked on the new chapter heading.[3] Scholars could search for terms relevant to their work across the whole collection or within subject groupings. The browser's *find* feature located words within a chapter. A pagination tool took the scholar to a known page.[4]

Every digital journals and books project has faced the challenge of deciding which features to include in a system design and under what conditions to modify that design. Substantial changes may confuse existing users of a collection. Such confusion may lead to misuse, resultant dissatisfaction, and a smaller chance that a scholar will return to the collection. However, if such modifications are true improvements, they should result in greater user satisfaction overall and more use in the long run.[5]

Context of the Online Books Evaluation Project

As anticipated, during the six years (1994–1999 inclusive) in which we planned and executed the project, relevant environmental elements changed substantially. Developments in the Internet and the World Wide Web, in com-

puter literacy and access within all levels of academia, in the development of digital libraries, and in scholarly publishing are summarized in the following sections.

National Environment—Expanded Potential Access to Online Resources

The national computing environment increasingly favored scholars' adopting computers, the Internet, and online resources. Prominent mainstream newspapers like *The New York Times* featured computers, the Internet, and related topics daily. The price/power relationship for personal computers improved enormously, with adequate computers available for under $1,000 since 1998. Penetration of personal computers and use of the Internet grew throughout American society. In 1999, over half of households owned at least one computer; and over 100 million American adults were using the Internet, up one-third from 1998. Users of email and the Web have become heavier and more sophisticated consumers.[6] But after 1997, typical consumers saw little gain in available speed of access to the Internet (56K at best) or in prices for ISP accounts.

The Columbia Environment—Increased Access to and Use of Computing and the Web

At Columbia University, even as it continued to grow, the modem pool for non-network access ran at near peak capacity, leaving scholars often frustrated in their attempts to dial-in to the campus network for email or electronic resources.[7] By 1999 most students owned computers. Students living in Columbia dormitories had Ethernet connections and Web access superior to that of the typical graduate student or faculty member working at home.

Scholars and Libraries—A Potentially Frustrating Relationship

As publication of journals and books expands, scholars face increasing difficulty in locating material that will be useful to their course preparation and research. At the same time, libraries struggle with the challenges of acquiring needed books and journals and making them available.

Scholars are frustrated by the limitations of the print library. In our surveys and interviews, they reported that they find that:[8]

- their libraries have not acquired the books they want to use;
- their libraries have acquired those books but they are not yet on the shelf;
- those books have entered circulation and are not now on the shelf and that the recall process will take many days;
- the online catalog lists those books as on the shelf but no one can find them;
- those books are on reserve and, hence, available at best for a period of time too brief to allow extended reading and review;

- browsing through the stacks that might have books relevant to their work takes substantial time and can only be done during the library's open hours;
- the library catalog includes too little information to indicate whether a book will be useful for a given research project.

Scholars are hopeful that online books will solve these problems.

Continuing Roll-Out and Expanded Use of Electronic Scholarly Resources

Coincident with the growing penetration of the Internet and development of the Web was the growth of digital scholarly resources. Digital abstracting and indexing resources are so popular that publishers now produce few paper versions. Electronic journals are increasing in availability and popularity over time. As time passes, more scholars are aware of what is available in electronic form, find these growing collections of value, and chose them.[9] Their familiarity with electronic versions of these two key categories of scholarly library resources leads scholars both to recommend them to their colleagues and to greater interest in all types of electronic resources. In the first seven months of 1999, Columbia scholars had 37 percent more accesses of the growing JSTOR collection of journal backfiles than they had in all of 1998. In the peak month, April, JSTOR use was equivalent to about one access per Columbia community member.

Findings of the Online Books Evaluation Project

Economics of Scholarly Book Publishing and Online Books

A key facet of this project was analysis of the lifecycle costs of scholarly books in print and online format. The online book format has developed slowly; as a result the industry of online publishing and the terms of providing online books are still in their infancy. With technology evolving as well, the various systems for and costs of publishing, distributing, maintaining, and owning online books are at the beginning of their developmental cycles. And it is not common to measure lifecycle costs for print books. But the available evidence permits order-of-magnitude comparisons.

Integrating Online Books into Publishing Process

In the future, production of online book files should be part of the publishing process, somewhere between the author's creation of a manuscript and the typesetter's creation of the film for offset printing.[10] Publishers or vendors will develop collections of online books, maintain them on central servers, and offer them as individual titles or groups of titles to libraries and individual scholars. Some publishers, e.g., Chadwyck-Healey, have begun to offer online texts in this way. NetLibrary is the first major intermediary vendor in the general library online book market.

As online books are not yet a standard publisher's output, our model assumed conversion of a print book, ASCII or other file format to HTML or SGML[11] as an add-on cost to the traditional publishing process.[12] Publishers or vendors also incur new costs in creating a permanent URL, maintaining a central server, and transmitting the online books to the reader via the Web or an alternative secured Internet mode. CD-ROM or DVD might be an alternative system for providing some electronic books, e.g., textbooks.

Print Books—Analysis of Costs to Scholarly Publishers

We analyzed pro forma revenues and costs for five print scholarly monographs from a major university press and estimates of the costs of producing and maintaining online versions of these books. These five books were expected to sell far better than most scholarly monographs—several thousand units each, while relatively few such books sell more than 500 copies. Thus, their production runs were larger and per unit costs lower than those for most scholarly monographs. In addition, these books were published by one of the larger university presses and, hence, enjoyed relative economies of scope in overhead costs. We analyzed the revenue and costs for these books in print form at length in a 1998 white paper on the economics of scholarly communication.[13]

Pro forma financial statements estimated that the publisher would sell a total of 13,650 copies (3,450 cloth copies and 10,200 paper copies) of these five books at an average net price of $17. The projected surplus (total net income minus total cost) for these books was $12,668, or $0.93 per copy sold.[14]

Online Books—Modest Incremental Costs to Publishers and Vendors

Hitherto, with few exceptions, online books have been produced after print versions were completed. Thus, the costs available for analysis are those of converting a print book to online format—generally either by scanning and further processing a print book or by taking an electronic file and converting it to SGML, HTML, or PDF.

Its Humanities Text Initiative has given the University of Michigan extensive experience in the first of these methods. The process of scanning, running the scanned text through optical character recognition software, encoding in SGML, and proof reading at all stages costs an average of about $1.51 per page. This does not include management of the project or the information systems.[15]

In mid-1998 Columbia contracted out HTML coding for books in this project. The coding was based on uniform standards set by Columbia staff. Coding for ASCII files cost $0.36 per 1,000 characters or an average of about $1.00 per page. Conversion from Quark format cost an average of about $2.15 per page. Then back at Columbia proof reading, fine-tuning and adding graphics to the files cost about $0.42 per page. Managing the conversion contract cost about $1,000 in staff time, or about $20 per book.

The University of Pennsylvania has found that when it receives clean PostScript files with fonts that its software can interpret easily, its system of converting to PDF costs four cents a page.[16] A student worker can create a final, web-ready PDF file of a well-behaved 300-page book in an hour. A troublesome book can take over five times as long, for a per page cost of twenty cents or more. These costs do not include overhead costs, but they are significantly lower than those for the two methods just described. This supports our theory that publishing can incorporate electronic formats for new books at a low incremental cost.

Columbia's staff calculated the full cost of maintaining books on a server as about $1.00 per MB per year. Books vary in size, but with some graphics a book might have about 44 pages per MB, for an annual cost of $0.023 per page per year. If a publisher were to maintain a book on a server for 30 years, the present value of the cost would be about $0.35 per page, or roughly $81 to $112 for each of the five books in our sample.[17] The cost of migrating these books over time might be in the same range.

The costs of the online version vary with the number of pages in the book as well as the method of conversion. They would also vary with the complexity of the books, i.e., the quantity of graphics, multimedia, and links to other online resources, but we have assumed uniformity in this analysis. For our five books, the present value of the lifecycle costs of production and maintenance might be approximately as follows:

	Base	With Migration	Break-Even Quantity
From Print to SGML	$2,677	$3,256	191
From ASCII to HTML	$2,628	$3,128	184
From Quark to HTML	$4,270	$4,760	280
From PostScript to PDF	$657	$1,157	68

It is much less costly to convert books from print or ASCII than from Quark. The PostScript to PDF method is the least costly at this time. However, we did not analyze the differences in the utility of the electronic books resulting from these methods. The weighted average net price per copy sold for these five books was projected at about $17. To cover the incremental costs of producing, maintaining, refreshing, and migrating the online versions of these books, the publisher would need to sell the above quantities of these books in online format at this average price.

This $17 net price is based on the mix of cloth cover and paperback copies that this press expected to sell. If the online books were sold at the cloth cover price (a weighted average of $30.46), the break-even quantity would be much smaller. On the other hand, if they were sold at the paperback price (a weighted average of $12.55), the break-even quantity would be much larger. These quantities are two-thirds or more of the total sales for many specialized mono-

graphs. However, they are a modest share of the projected sales for these five books.

The publisher or intermediary might charge the purchasing library a service fee for maintaining the book online, migrating it regularly, and the like, as netLibrary does. If so, these costs could be recouped over time rather than through initial sales.[18]

Online Books Have Lower Lifecycle Costs for Libraries

One of the project's implicit hypotheses was that online books would have lower lifecycle costs of ownership for a library. Two expectations underpinned this hypothesis: (1) Print books will require ever-more-expensive storage space and manpower for acquisition, processing, and circulation; and (2) Online books will require decreasingly costly computing hardware and little staff time.

We estimated the current present value of a library's lifecycle costs for both formats. Table 1 estimates the present value of the basic cost elements at Columbia for both types of books. Assuming a $50 purchase price for both formats, the present value of the total lifecycle stream of costs is about $156 for the print copy and $127 for the online book, for a 19 percent savings with the online version.

A scholarly library incurs costs for the Internet and computing infrastructure that allows it to provide electronic resources to its community.[20] However, institutions have incurred these costs already, even though their libraries have not provided online books. We have not included these sunk infrastructure costs in our calculations of the costs of providing individual online books. If an institution would assume other unique costs in providing its community with online books, an analysis of this sort should include those costs. One such cost might be that of expanding computer-printing capacity so that scholars could print out portions of online books. Universities and their libraries must decide the extent to which they will pass on these costs to their scholars.

Use and User Reactions

Use of the books in the online collection was measured via server logs, including data on individual users. Reactions were obtained from various surveys and interviews.

Online Reference Works Were More Used Than Their Print Counterparts; Format and Timeliness Were Critical

Our collection of six reference works saw varying patterns of use, but all were apparently used more often in online form than in print form.[21] Use of *Columbia Concise Encyclopedia* and *Columbia Grangers World of Poetry,* two older works not aimed at a university audience, declined substantially to a few hundred sessions per semester. Librarians reported that those resources are used only a few times each month in paper format. *The Oxford English Dictio-*

Table 1. Libraries' Lifecycle Costs of Book Ownership

	Print Book	Online Book
Purchase Price, Average	$50.00	$50.00
Selection	$3.59	$3.59
Processing:		
Ordering		$2.00
Locate & Handle Bibliographic Record		$5.92
Receive Physical Item		$0.00
Payment	$43.67	$2.00
Initial Physical Processing		$0.00
Cataloguing		$25.00
Storage	$4.61	$0.00
Average Cost of Circulation	$43.97	
Stack Maintenance	$5.47	$38.43
Collection Maintenance	$1.90	
Repair/Rebind	$0.28	$0.00
Replace—New Book & Processing	$2.08	$0.00
Total	**$155.57**	**$126.94**

Selection: Estimated from share of hours spent by librarians and assistants at relevant salary and fringe rates divided by number of new items. (Assumes average librarian salary of $45,000; staff assistant salary of $22,000; and student assistant wage of $8 per hour.) Most books are purchased via approval plans, not individually selected and ordered.

Storage: Present Value of 30 years at $0.30 per year with 5 percent interest rate.[19]

Cost of Circulation for Print: Present Value of 30 years with 5 percent interest rate at $2.86 per circulation and an average of one circulation per year.

Stack Maintenance: Includes shelfreading, shifting.

Collection Maintenance: Includes searching for and tracking missing books.

Replace Print Book: 2,500 volumes lost and 116,000 purchased annually: 2.16 percent loss rate, assumed value replaced over 30 year period in purchase price and processing.

The online model assumes that books will be bought via a system with terms negotiated with one or several publishers or intermediaries, e.g., approval plan, user selection and mass availability, user selection and individual availability, etc. Costs of ordering the books will be similar to those via approval plans for print books. Cataloging costs are estimated at the original cataloging level experienced in the Columbia experiment, rather than at the much lower copy cataloging level that would prevail in the long run. (Alternatively, publishers might take on the cost of cataloging online books.)

nary was the most used resource in the collection; 1,370 unique scholars executed almost 29,000 hits on the Web version in spring 1999. Use of *African-American Women, Native American Women,* and *Chaucer Name Dictionary* fluctuated from semester to semester, always in the several hundred hits or sessions range.[22]

Monographs and Humanities Texts Had Modest but Growing Use

Use of the three Chadwyck-Healey humanities full-text databases (*English Poetry Database, English Verse Drama, Patrologia Latina*) grew over time, but

remained relatively modest. During spring 1999, each had from 52 to 122 users.

From 1996 to 1998, use of the 54 *Past Masters* classic texts declined by one half to 3,384 hits. However, the number of hits in spring 1999 was 58 percent greater than a year earlier. Use of these texts was largely concentrated in a small share of the titles that were used in classes in political philosophy and theory. In the period July 1996 to June 1999, one-seventh of the texts received two-thirds of the total hits.

The online books collection included 36 monographic titles as of July 1997, 55 as of July 1998, 68 as of year-end 1998, and 108 as of June 1999. Twenty-five of these books were assigned in one or more courses for one or more semesters during this period.[23] These monographic books received 3,542 hits in 1997, 4,885 hits in 1998, and 2,919 hits in the first half of 1999, a roughly linear steady increase. In spring 1999, there were 806 cases in which an individual scholar used one of these titles one or more times. As some individuals used more than one online book, the number of individual users was smaller.

Books were Used Significantly in the Online Format

Books that were available online may have been used by more scholars in online format than in paper format. In spring 1999, nearly three times as many scholars "clicked" on the average online monographic book as "circulated" its print version. [24] Using the so-called *Principle of Use Until Satisfaction* (Kantor, op. cit.), we assumed that any encounter between a scholar and a book was equally likely to represent a complete use event. That is, while some encounters were longer than others, in each case the user could continue until satisfied. In the use of a traditional paper book, that use could have been as little as a quick look at its Table of Contents while standing at the stacks or as much as checking it out and reading the entire book. The average number of hits per monographic online book per user, per half year, hovered at four to five throughout the study period. Thus, on average, the use of the online books had depth beyond clicking on and looking at the Title Page-Table of Contents file.

Only one-fifth of the users of monographic online books used more than one or two titles throughout the study period. This is not surprising, given that we lacked the critical mass of different books that would attract a scholar repeatedly.

Online Book Use Occurred in a Mixed Format Environment

Surveys and interviews indicated that scholars were not generally "reading" books online. They tended to browse the books online and then to print out relevant portions or to look for print copies for extended reading. Sometimes they referred to the online version to track down a quotation or a citation. Other studies have found that this pattern is common in using electronic journals as well.[25]

Online Book Use Varied by University Cohort

The distribution of use among Columbia sub-communities varied by book category. For some books, undergraduates were the primary users, for others graduate students in the Arts and Sciences or in professional programs were dominant. Faculty members were never more than a few percent of users.

Online Books did Not Displace Personal or Library Copies

An in-class survey, administered when part of a book in the online collection was required reading in a course, revealed that a growing share of students used the online versions in some way,[26] but few "read" online. By spring 1999, 39 percent of the students in these classes indicated that some form of an online book was used to read the assignment, the same share used their own copies of the book. The student's own copy of a book remained both the most common single method of reading an assignment and the single most preferred method.[27] However, at that point, 43 percent preferred some form of online book use while only one third preferred to use their own copies.

Cost and the Use of Books

Supposing for the moment that a scholar is familiar with online books, has easy access to them, and incurs no cost for use of an online book, we can illustrate how scholars seem to feel about these books with a simple preference table. There are many modes of using a book, but for simplicity we divide them into two: *read much of the book* and *read little of the book.* Similarly, there are many ways that we can describe the cost of a paper version, but we simply divide them into two cases: *low* and *high.* If a scholar intends to read little of a book, the ability to locate things within it will be more important, while comfort of reading, annotation, etc., will be less important. In this case an online book will be preferred, without regard to the price of the paper version. However, if the scholar intends to read much of the book, the inconvenience of online reading will be the dominant factor, forcing the scholar into a *buy versus borrow* decision. The following simple table shows these four cases.

Scholar's Preferences for Book Access

	Read Much	Read Little
Low Cost Book	Buy	Online
High Cost Book	Borrow	Online

This table makes it clear that, if a library wishes to respond to its scholars' preferences, its collection of paper books should continue to exist to serve the substantial reading, high cost situation.[28] Two of the quadrants call for the library to provide online access to electronic books. As the readability of online books improves, and scholars' habits evolve, the meaning of *read much* will

change, so that the two quadrants on the right of the table will steadily expand as a share of total preference for use of scholarly books. If the library is to remain a key resource for scholarly access to monographs, its online holdings must expand to keep pace with these evolving preferences.

Librarians Find Appeal in Online Books

In interviews and other discussions from summer 1998 to fall 1999, college and university librarians expressed great interest in and considerable optimism about the potential for online books in their collections. They viewed reference works as having particularly great utility in online form. But they also saw value in having books that are in high demand, of transient topicality, or not part of their print collections available in this format. Fewer were willing to contemplate acquiring online versions of books that are single copy components of their standard research collections, mainly because they could not afford to pay for such a book twice.

Librarians were concerned about how the marketplace for scholarly books will evolve and particularly how online books will be provided. How will online books be packaged and priced? What guarantees of availability in the short and long run, preservation, format updating and the like will publishers or vendors provide? What conditions of use will publishers and vendors seek? What will happen to the important copyright concept of *fair use*? Librarians also hoped to see cataloging and usage statistics provided by vendors, as well as user-friendly design of both the general interface to online books and the individual books.

Potential Market Arrangements

Our findings on scholars' interests in online books and on the costs of providing online books suggest that scholarly publishers should test new models that combine print and online availability of books. With the right mix of offerings, the overall market should expand. The goal of these models would be to increase availability of books to scholars and profitability to scholarly publishers. Without market experimentation, we cannot know which combination will most enhance the social good.[29]

Mixed product models could expand research library markets, develop new library markets, and attract more individual scholars as buyers. Possible options include:[30]

- Retain and expand the research library market with:
  - — A free online version, i.e., an enhanced product package, or an online version available at a small incremental price, when the library purchases the print version.
  - — Modestly priced online collections offered to libraries that buy a

substantial share of the titles in print version. This would likely expand sales of the print copies somewhat.

- Expand sale of scholarly monographs into the libraries in the U.S. and abroad that purchase few such books. Price online collections for sale to consortia. Make an on-demand electronic version available as an attractive alternative to interlibrary loan.
- Encourage individual scholar's purchases of monographs with:
  - — Pay per electronic view. This could be attractive to scholars not having library access to print or electronic copies.
  - — Online ordering of a print copy at a discount from the site of a library's online version.
  - — Lowered prices for the print copies to reflect their modest marginal cost.

The processes of designing, producing, and marketing online books are in their infancy. Some scholars who have begun to author books with an online format in mind find it much more challenging than anticipated, raising questions such as: How should primary documents be included? How should they be incorporated into the analysis? Will the text automatically update itself? Such issues do not arise in authoring traditional scholarly monographs.

Over the next several years the concept of online books will resolve itself so that scholars, authors, publishers, vendors, and libraries will be able to determine what combinations of modes and models will be most beneficial for the scholarly community as a whole, including all of these stakeholders and the readers as well. In the final analysis the entire system is one by which authors communicate to readers across spans of both space and time. The present system has evolved in a way that supports value-adding transformations at several points: the publisher adds quality and authentication; the vendor adds distributional services; the library adds organization and preservation, as well as further quality certification. With electronic books there is no reason to sacrifice any of these values. However, the final economic driver will be the value delivered to readers, which is the ultimate quid pro quo that leads to the injection of revenue into the process. As technology permits us to blend several of the traditional stages of the communication chain, ultimately economic models will more accurately reflect that value delivered to readers, and will correspond more closely to their allocation of attention to the works, whether in reading or in citation. Writers, editors, and publishers of online magazines assess specific articles by logging the amount of attention that readers pay to them. The scholarly world is also very much concerned with the capture of attention, as measured by citation analysis. Thus it is likely that such analyses will be introduced at several points in the chain of communication, and will, eventually, become the basis for assignment of costs and prices.

Notes

1. Many projects working with books that were already published have used a system of scanning a book's pages and using OCR to create an index of its contents.
2. The project was not able to test enhancements as we converted existing books rather than creating new products. Other systems such as PDF now have greater capabilities than in 1995. Publishers would be wise to evaluate the trade-offs of the various book production and display options.
3. Web enhancements over the five years of the project allowed us to add features. However, we could not economically modify the books already in the collection. Our designers decided that consistency within the collection was more important than adding features for only the more recent books.
4. Page numbers were not used within the text, so scholars could not cite a page number in a reference to the online book. Scholars requested pagination equivalent to that in the print books.
5. See, for example, Bishop et al (2000). This reports on the NSF/DARPA/NASA Digital Libraries Initiative project at the University of Illinois.
6. Estimates by the Strategis Group reported in *The New York Times on the Web*, November 10, 1999. http://www.nytimes.com/library/tech/99/11/biztech/articles/10net.html.
7. Many Columbia scholars residing off-campus had ISP accounts as well as the ability to dial-in to Columbia directly. The online books were accessed via a sign-in with one's Columbia ID, so that a scholar could reach them using any Internet route.
8. Kantor (1984) discussed analysis of some of these problems.
9. Electronic journals projects have found that the breadth and depth of a collection are critical to user adoption. In addition, it takes time for scholars to become aware of the availability of these resources and to find an opportunity to use them. See Bishop et al (2000) and *SuperJournal* (1999).
10. This last step might be eliminated for a class of low demand books for which desktop publishing and on-demand printing might become standard.
11. If a book were marked up in SGML, the standard is to convert it to HTML for Web viewing *on the fly*.
12. For its current online books project the University of Pennsylvania is using a system that involves distilling PostScript files into PDF files and reassembling the chapters into a single file. Compose, a plugin, builds bookmarks and links from the book index to the pages. (Roy Heinz, Director, Information Systems, University of Pennsylvania Library, provided information on their process for conversion and the costs that they are experiencing, personal communication.)
13. Summerfield (1998).
14. This press, like many university presses, receives free space and utilities from its parent university so this analysis does not reflect all economic costs of this enterprise.
15. Christina Powell, Coordinator, Humanities Text Initiative, personal communication.
16. Roy Heinz, personal communication.
17. This present value estimate assumes a real cost of money of 5 percent per year as well as a 30 year time span. These parameters are used in the later estimates of the lifecycle costs of a book for a library as well.
18. The vendor could charge higher rates in the early years than support actual costs (knowing that the book had greater value to the libraries while it was relatively new) and use the excess funds to fund those activities in the later years. The vendor could promise to maintain books that become stale quickly for only a few years and then take them down when few libraries were willing to support them. Of course, this thwarts libraries' archival mission.
19. Getz (1993). This is the cost of off-site storage of a single volume.
20. A college or university uses these resources for many purposes beyond library services.
21. We lacked a firm count of use of reference resources that were kept on open shelves as scholars could use and reshelve them at will. Librarians could track use of books kept behind their service center.
22. These resources were available in two formats for which the measuring systems varied. It is impossible to tell for certain if declines in use of the text-based CNet versions, measured in sessions, were offset by gains in use of the browser-accessible CWeb versions, measured in hits.
23. This count is based on faculty having put a book on reserve for a course. We had no way of determining if an instructor used a book in a course but did not put it on reserve.
24. We could not track use of print-on-paper books that did not include checking them out so we lack a total count of use of books in that format. Libraries typically estimate that in-library use of books equals or exceeds use through circulation.
25. See Bishop et al (2000) and *SuperJournal* (1999).
26. Forms of using an online book included reading online, printing out the online version, saving to disk.
27. The responses to these questions would depend on the nature of the assignment. If only one chapter of a book was required reading for a course, our understanding is that few students would be likely to purchase that book. On the other hand, if the whole book was assigned and the book was available at a reasonable price, a much greater share of the students would be likely to purchase it.
28. Of course, textbooks fall into this category and libraries serve only a small share of their communities' demands for textbooks.

29. The University of Michigan PEAK project of journal pricing was one such experiment.
30. These market models are discussed in Mandel and Summerfield (1998).

References

Ann Peterson Bishop, Laura J. Neumann, Susan Leigh Star, Cecelia Merkel, Emily Ignacio, and Robert J. Sandusky. "Digital Libraries: Situating Use in Changing Information Infrastructure." Preprint of Journal of the American Society for Information Science 51, no. 4 (March 2000).

Getz, Malcolm. "Information Storage." *Encyclopedia of Library and Information Science* 52, supplement 15 (1993), 201–39.

Kantor, Paul. *Objective Performance Measures for Academic and Research Libraries*. Washington, D.C.: Association of Research Libraries, 1984.

Mandel, Carol and Mary Summerfield. "Scholarly Monographs Online: Potentialities and Realities Suggested by the Columbia University Online Books Evaluation Project." *http://www.arl.org/scomm/epub/papers/mandel.html*. January 1998.

"Summary of SuperJournal Findings: Readers." *http://www.superjournal.ac.uk/sj/findread.htm*. Draft April 26, 1999.

Summerfield, Mary. "Issues in the Economics of Scholarly Communication." *http://www.columbia.edu/cu/libraries/digital/texts/econpap.html*. March 1998.

Summerfield, Mary and Paul Kantor. "Online Books Evaluation Project: Analytical Principles & Design." *http://www.columbia.edu/cu/libraries/digital/olbdocs/protocol/*. May 1996.

Summerfield, Mary, Carol Mandel and Paul Kantor. "The Online Books Project, Columbia University: Final Report." *http://www.columbia.edu/cu/libraries/digital/olbdocs/finalreport.html*. December 1999.