

The Value of Purchasing E-book Collections from a Large Publisher

Aaron K. Shrimplin

Miami University of Ohio, shrimpak@muohio.edu

Jennifer W. Bazeley

Miami University of Ohio, bazelejw@miamioh.edu

Follow this and additional works at: <http://docs.lib.purdue.edu/charleston>

An indexed, print copy of the Proceedings is also available for purchase at: <http://www.thepress.purdue.edu/series/charleston>.

You may also be interested in the new series, Charleston Insights in Library, Archival, and Information Sciences. Find out more at: <http://www.thepress.purdue.edu/series/charleston-insights-library-archival-and-information-sciences>.

Aaron K. Shrimplin and Jennifer W. Bazeley, "The Value of Purchasing E-book Collections from a Large Publisher" (2011). *Proceedings of the Charleston Library Conference*.

<http://dx.doi.org/10.5703/1288284314926>

This document has been made available through Purdue e-Pubs, a service of the Purdue University Libraries. Please contact epubs@purdue.edu for additional information.

The Value of Purchasing E-book Collections from a Large Publisher

Aaron K. Shrimplin, Assistant Dean, Collections and Research Services, Miami University Libraries
Jennifer W. Bazeley, E-Resources & Serials Librarian, Miami University Libraries

Background and Context

Miami University is a public university with an undergraduate student population of 14,936. It is known for its teaching and learning and has a relatively small graduate student population of 2,459. Miami University Libraries, like many academic libraries, are transitioning to a mostly digital library. We are in the final phases of transitioning almost all of our journal holdings to e-journals and are in the early phases of shifting our monograph collections to e-books. To date, we have been experimenting with a variety of providers, platforms, pricing, and selection models. Most of our e-book acquisitions have been done in partnership with our consortium, OhioLINK, and most of these deals have been publisher deals. In addition, we have been piloting a Patron Driven Acquisition (PDA) model with ebrary and YBP, are exploring the option of Short-Term-Loans (STL), and librarians are beginning to privilege e-books over print books as they develop their subject areas.

User Attitudes toward E-books at Miami

Two key studies have been completed at Miami University to determine campus attitudes toward e-books. The first study used Q methodology to unveil four distinct viewpoints toward e-books: Book Lovers, Technophiles, Pragmatists, and Printers (Shrimplin, 2011). The second study used a large-n survey (n=735) to determine what percentage of the campus falls into each of the four viewpoints (Revelle, 2012). It was determined that 34% are Book Lovers, 23% are Technophiles, 17% are Pragmatists, and 26% are Printers. In a nutshell, Book Lovers have an inherent affinity for the print form. Technophiles are strongly interested in the possibilities of new technology as regards the book. Pragmatists are the most neutral of the four, as they are most interested in content and see pros and cons of both formats. Printers prefer print books but are distinguished from Book Lovers in that they have specific difficulties with usability of e-books.

Where Do We Go from Here?

It's clear our next steps involve ramping up our e-book collections, both locally and through the consortium. What isn't so clear is how we do that – what is the best strategy for building e-book collections? At this point-in-time, we have more questions than answers. Should we lease content or own it? Is a local load into OhioLINK's Electronic Book Center (EBC) necessary or can we exclusively access the content from the vendor or publisher website? Do we acquire e-books at the collection level or title-by-title? How far should we extend the PDA model? Given all these questions, it seems appropriate to spend some time trying to assess some of our experiments to-date so that we have better clarity about how to move forward with scaling up our e-book collections. What does the usage data tell us? More specifically, what's the value of purchasing e-book collections from a large publisher? To answer this question we initiated a preliminary study that focuses on the 2008 Springer e-book collection (purchased through OhioLINK) and its use over a three-year time period.

Springer E-book Collections and SpringerLink

Springer e-books are divided into twelve subject collections. More than 48,000 e-books are currently available in the Springer e-book collections. They have no DRM, and once purchased, the owner has perpetual access. The collection is strongest in STM titles, but there are humanities and social sciences titles. Types of books include monographs, proceedings, contributed volumes, textbooks, and reference books. Each book chapter is a downloadable PDF on the Springer site. Springer e-journals are also accessed at this site. Springer provides COUNTER compliant usage reports for all materials accessed at SpringerLink.

OhioLINK's Electronic Book Center

Miami University Libraries, through OhioLINK, has purchased to-date 25,571 Springer e-books. In addition to the SpringerLink platform, Miami users

have the option to access Springer e-book titles on OhioLINK's Electronic Book Center (EBC). OhioLINK's EBC is an electronic book platform created and maintained by the OhioLINK consortium. Electronic books purchased by the consortium (with the exception of EBSCO E-Books) are loaded into the EBC. OhioLINK loads e-books into their EBC both for access and preservation purposes. Some e-books can be viewed both in the EBC and at the publisher's platform. Purchased e-books include scholarly and reference works from the following publishers: ABC-Clio, Chadwyck-Healy (ProQuest), Gale, Oxford University Press, Sage, and Springer. In addition to hosting these locally loaded e-books, the EBC also contains bibliographic information and links to e-books purchased (EBSCO E-Books) and leased (Safari Books Online, Oxford Reference Online, and ACLS Humanities E-Books) by the consortium and hosted only on the publisher platform. Springer e-books can be viewed in the EBC as HTML, or individual chapters can be downloaded as PDFs. OhioLINK provides usage reports for these titles in the EBC, but they are not COUNTER compliant.

Methodology

One of our primary goals in conducting this study was to find out how Springer e-books are getting used to better understand the model and its costs as we go forward with e-book acquisitions. Because usage data reside on two different platforms, it was not a simple process to compile usage statistics. The COUNTER compliant data we obtained from Springer were easy to use, and we merged this usage data with Springer's e-book Title List. Merging the two data sources together provided us with additional title level information, such as copyright year, subject information, and book type (e.g., reference work, monograph, textbook, etc.). The usage data from OhioLINK's website didn't have the granularity of the statistics provided by Springer. For example, we were not able to produce a list of titles that had not been used for a specific period of time. OhioLINK's e-book usage tool lets you pick a collection, select a date range and limit to a member library. The results are provided as an HTML page and provide the title and total number of downloads. Our next challenge was to merge our Springer data that included title-level information with the usage

data created by the OhioLINK E-book Usage Reporter. For this task we used SPSS, and by matching on title, we were able to merge the two data files together. Our analysis of the collection focused on books with a 2008 copyright year and that were added to the EBC in 2008. For these titles (n=2,529), we looked at usage over a three-year time frame (2008-2010).

Findings

Title Use

Of the 2,529 titles in our sample, 23% had at least one download over the three-year time frame. Put differently, 77% of the sample titles had no use from 2008-2010. How does this finding measure up with other studies? Gavin and Kent found that 40% of print books are unused six years after purchase (Gavin, 1977). Looking at our data, an average of 194 titles are accessed for the first time each year. If we were to use this average and extend it out over a six-year time frame, we calculate that 54% of our sample titles will be unused after six years. Bucknell found fewer than 40% of the purchased Springer e-books at his library have not been used within two years of purchase (Bucknell, 2010). While Sprague found that approximately 19% of e-book titles from NetLibrary and ebrary (43,723) have been accessed at least once (Sprague, 2009).

Pareto Principle

Our study found that a small number of titles are heavily used. In fact, we confirmed the Pareto Principle or the "80/20 Rule" when we look at titles used and total downloads. That is to say, 20% of the used titles (116/582) account for 80% of total downloads (6,370/7,963).

Long Tail

Our findings show that a few high use titles dominate. The number one most used title, *Algorithms and Data Structures*, had 2,264 uses, or 28.4%, of the total uses over the three-year time period. Our findings also show a long tail of infrequently used titles. Analyzing the data, we find that of the 582 titles in our sample that were used at least once over the three year time frame, 318 had three uses or less (54.6%). Furthermore, these 318

titles constitute 6.2% of the total downloads (7,963).

Inflationary Effect

Local analysis of e-journal usage has found that in recent years roughly half of all downloads have taken place on vendor or publisher websites. In other words, half of the downloads of e-journal articles by Miami University patrons are done on publisher or vendor websites rather than from OhioLINK's Electronic Journal Center (EJC). This suggests that a sizable number of users are using search engines (e.g., Google Scholar) to find journal articles or they are going directly to a publisher's website for content. When we examine Springer e-journal usage, in particular, we find support for the overall trend of increasing usage on publisher websites. Because SpringerLink cross-references books and journals, we speculate that as usage on the journal-side of SpringerLink increases, usage for e-books on SpringerLink should also increase. In short, an inflationary effect would result when users were able to search and link to journal and book content. This cross-linking is not currently supported on the OhioLINK platform. As Figures 1 and 2 show, as downloads increase on SpringerLink, we also see an increase year-over-year in the number of SpringerLink e-book downloads.

Downloads by Book Type

Looking at book type, the data suggest that textbooks are especially popular. While they are a relatively small percentage of the titles in the collection, they have a very high percentage of downloads. Figure 3 presents the percentage of titles by book type and the percentage of downloads by percentage of titles. It is clear that textbooks, and to a lesser extent reference materials and professional books, are used heavily relative to their percentage of titles in the collection. Figure 4 uses the same data as Figure 3 but presents it visually in the form of a word cloud.

Number of Titles Used by Subject Area

When we look at titles used by subject area, two findings stand out. First, four subject areas are in high demand: Computer Science, Engineering, Medicine, and Biomedical and Life Sciences. In short, Miami University users look for STEM con-

tent from Springer. Second, it appears that in almost all subject areas the most titles are used during the middle year of our study, 2009.

Past Usage Predicts Future Usage

Usage pattern shows that many titles that were well used (four or more) in 2008 received additional usage in 2009 and 2010. There are 87 titles with four or more uses in 2008. Of these 87 titles, 35 or 41.2% had one or more uses in 2009, and 20 or 23.0% had one or more uses in 2010. Figure 5 presents data showing heavily used titles (four or more downloads) during 2008 and suggests that these heavily used titles also tend to show additional usage in follow-up years. Each rectangle represents a book (87 in total). The size of the rectangle represents the number of downloads in 2008 and brighter green indicates more downloads during 2009 – 2010.

Value of the Publisher Deal

Our price tag to OhioLINK is a good value and becomes a better value over time with additional usage. Cost per download also seems to be a good value when compared with others (e.g., an Elsevier study of ScienceDirect e-books determined a cost per use of \$5.10 per chapter). Having the choice to select only a subset of subjects might be a useful option given our findings. Given the data and our findings, an ideal acquisition and pricing model would take into consideration the 80/20 finding and the long tail. Ideally, we would only want to buy the 20% of the titles (116) that had 80% of the use. As for the long tail or the 318 titles with 494 downloads, we would want to rent these titles and pay a cost per download charge. This model, based on the usage patterns discovered in this study, would maximize access and minimize costs.

Conclusions

Are e-book collections an underperforming asset? Our analysis of a sample of Springer e-books finds 77% of the titles were not used from 2008-2010. Of those titles that were used, the 80/20 Rule applies: 20% of the used titles account for 80% of all downloads. As a consequence, a few high use titles dominate usage and the long tail accounts for a very small percentage of the total downloads. Another finding suggests that platform does mat-

ter. E-books that are cross-searchable with e-journals (as is the case on SpringerLink) correlate with higher use. Given what we know about Miami University attitudes toward e-books, a usable, robust platform with book and journal content is very appealing to Pragmatists and Technophiles. Given our campus make-up as a primarily undergraduate university with an emphasis on teaching and learning, it makes sense that textbooks are in high demand. We also find that title use peaks across all subject area in 2009 and that there is a strong interest in STEM titles with Computer Science as the most used collection. Lastly, heavy use appears to beget more use as time goes on.

Next Steps

Going forward, we plan to continue to evaluate different e-book purchasing and pricing models with an understanding that no single approach will be adequate. Given that the use of the Springer e-book collection was concentrated in a relatively few titles, we will continue to experiment with patron-driven models. Rentals, short-term loans, and pay-per-use models are also options going forward. Comprehensive searching is useful to users and easy access and better usability should be criteria by which we consider future e-book collection purchases. We also have plans to continue to investigate the Springer e-book collections with more years and more titles. Doing a comparison with a

Humanities and Social Science collection, such as Oxford University Press (OUP) is also in the cards. As we move forward to ramp up our e-book collections we will continue to look for solutions that provide even more e-book titles, available at the time of publication, with multiple purchasing and pricing models.

References

- Bucknell, T. (2010). The 'Big Deal' Approach to Acquiring E-Books: A Usage-Based Study. *Serials, 23*(2), 126-134.
- Gavin, T. J. (1977). Use of a University Library Collection: a Progress Report on a Pittsburgh Study. *Library Journal, 102*(20), 2317-20.
- Revelle, A. K. (2012, July). Book Lovers, Technophiles, Pragmatists, and Printers: The Social and Demographic Structure of User Attitudes Toward E-books. *College & Research Libraries, July 2012*.
- Shrimplin, A. K. (2011, March). Contradictions and Consensus--Clusters of Opinions on E-Books. *College & Research Libraries, 72*(2), 181-90.
- Sprague, N. a. (2009). Assessing E-Books: Taking a Closer Look at E-Book Statistics. *Library Collections, Acquisitions, & Technical Services, 32*, 150-57.

Figure 1 Inflationary Effect: E-Journals

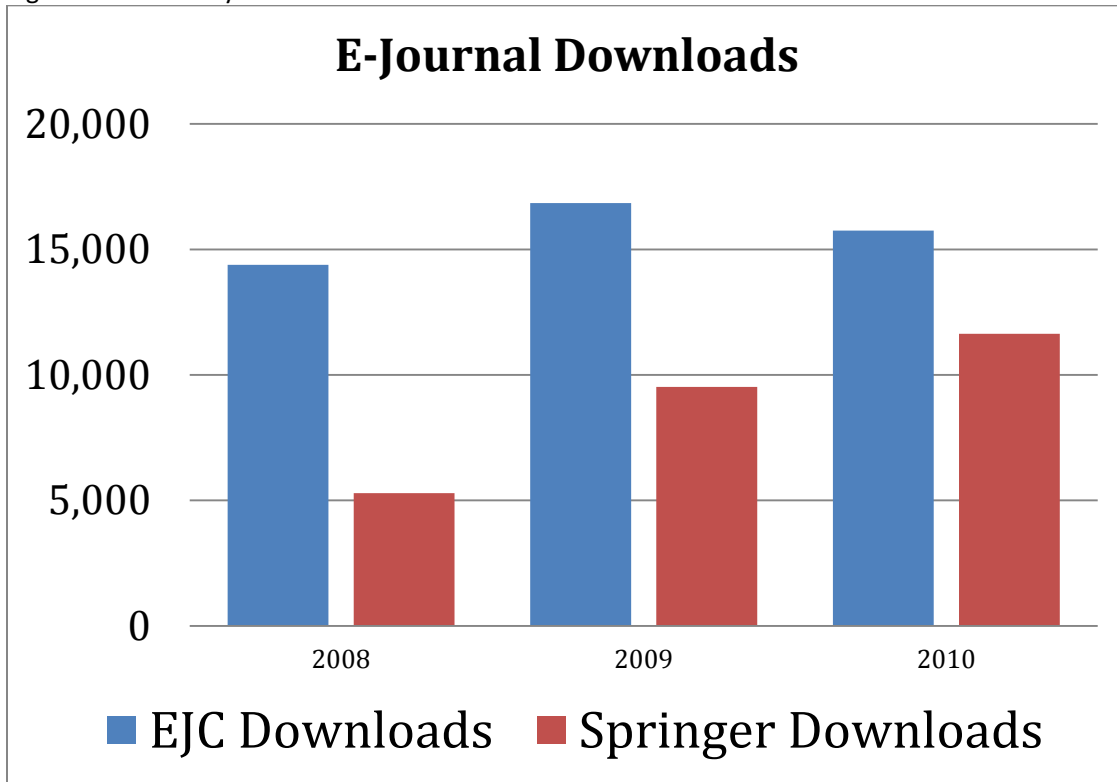


Figure 2 Inflationary Effect: E-books

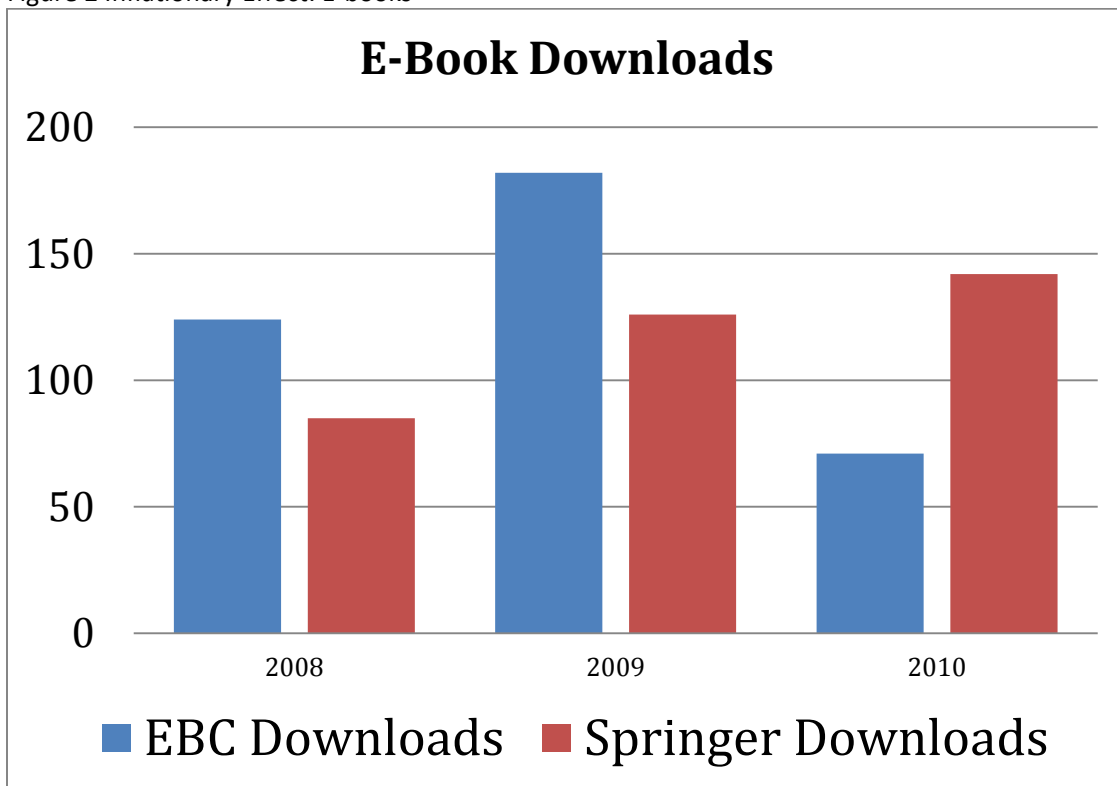


Figure 3 Number of Downloads by Book Type

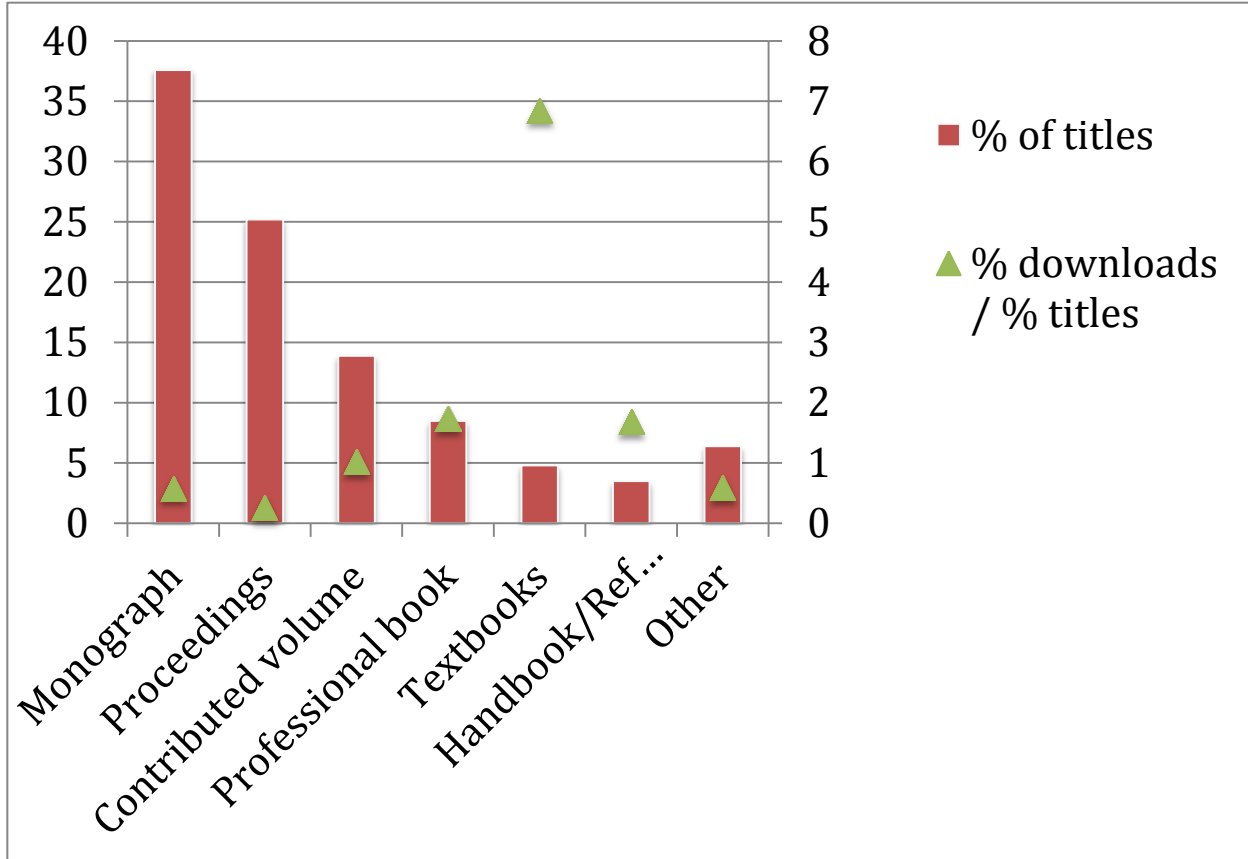


Figure 4 Book Type Word Cloud

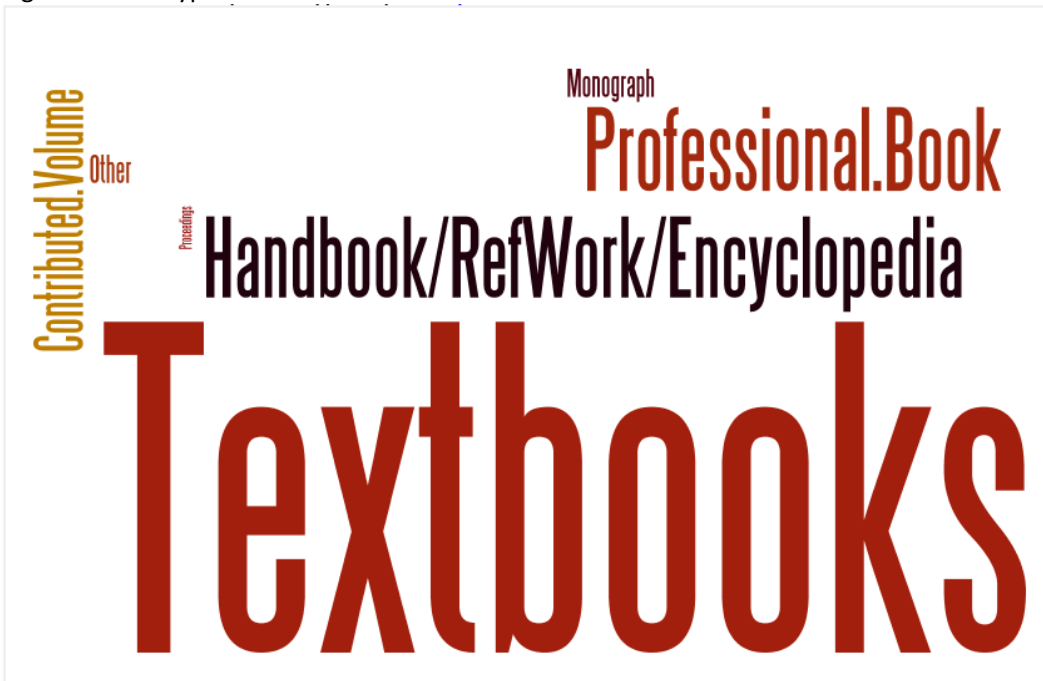


Figure 5 Past Usage Predicts Future Usage

