Comparing DDA E-Book Program Variances of Eight Large Academic Libraries

Kay Downey
Kent State University, mdowney1@kent.edu

Yin Zhang
Kent State University, yzhang4@kent.edu

Kay Downey and Yin Zhang, "Comparing DDA E-Book Program Variances of Eight Large Academic Libraries" (2017). *Proceedings of the Charleston Library Conference.*
http://dx.doi.org/10.5703/1288284316702

Comparing DDA E-Book Program Variances of Eight Large Academic Libraries

Kay Downey, Kent State University

Yin Zhang, Kent State University

Abstract

Researchers at Kent State University Libraries have collected DDA (demand-driven acquisition) e-book program data from eight large academic libraries in order to further research on DDA use and efficacy. As libraries transition more and more to the just-in-time acquisition model, it is even more imperative to understand the factors that contribute to successful collection management practices and sustainability. With multiple years of data from eight institutions, this will be the first large-scale study of this kind. In this study DDA e-book data was examined from each institution that detailed order and license information, bibliographic data, and usage data for each program. All institutions used ProQuest as their e-book provider and YBP as the DDA e-book jobber. A survey was also used to gather information about the parameters of each DDA discovery pool program variables. Formula logic based on actual usage data was developed in order to have an apples-to-apples comparison of overall cost under various DDA model scenarios. This study offers perspectives and considerations for implementing and evaluating a DDA program in large academic libraries. It also analyzes DDA e-book program commonalities and future directions that help librarians in choosing the e-book DDA business model that works best for their library.

Kent State University Libraries (KSUL) has been using the DDA e-book purchasing model since January 2012. Over the course of the DDA program two Kent State University professors, Yin Zhang and Kay Downey, have been conducting studies that take a close look at DDA business models. Studies focus on examination of the DDA variables that contribute to the best value and were the best fit for the Kent State University Libraries. Several years of accumulated DDA data has made it possible to conduct a systematic comparison of different demand-driven acquisition business models based on e-book usage patterns in order to figure out which model is most cost effective. The purpose of the current study is to determine if the conclusions drawn from the Kent data apply to other university libraries.

To perform the study, permission was obtained from seven other large academic libraries to acquire data about their DDA e-book programs. The content of the presentation included an overview of the DDA model, data collection, a description of DDA program commonalities and variances among the study participants, longitudinal analysis of e-book usage, DDA business model scenario analysis, and concluding findings and observations. Two common DDA business models were analyzed, the DDA model with the short-term loan (STL) component and the straight DDA with no short-term loans. The short-term loan model is structured to trigger an e-book purchase only after a set number of short-term loans have been triggered. The final cost in the STL model is the sum of each rental fee plus the list price of the e-book. Comparison and analysis of this data helped discover similarities and differences between the programs.

Analysis of the data raised a number of questions and issues, such as which report(s) would be most helpful, which program variables should be taken into consideration, what were the major differences and commonalities between programs, how do publisher conventions influence return on investment outcomes and how were these components correlated to determine which DDA business model is most effective.

One important note regarding the trigger definition involves the ProQuest migration from the ebrary platform to the new Ebook Central platform. The new platform redefines the trigger threshold to five minutes of use and one trigger, copy, print, or download. For this study the 10-10-1-1-1 trigger was used because that is the paradigm associated with the historical data used in this study. Another important concept is the definition of a session. One session is defined as the number of times a title is opened and incurs any use. A number of reports were available, but for the most part the Title report and the Trigger reports were most useful. Common to all of the datasets was the unique ebrary ID number. Purchase, bibliographic, and usage data were also obtained from the combined reports.

https://doi.org/10.5703/1288284316702

The following datasets were used: an e-mail survey of study participants; the ebrary Title report and Trigger report, which provided usage data, details of DDA transactions, license and cost details, publisher information, and the trigger event. Also used was the ebrary documentation, which provided current license restrictions and short-term loan rates for all publishers in the sample.

In addition to data reports a survey was issued with the hope that it might shed some light on some of the data inconsistencies that we had observed. The survey included questions about weeding, publisher exclusions, price caps, and general comments about their program. Survey results revealed that some institutions started their program as early as 2009 and some as late as 2014. Institutions 1 through 4 used the straight DDA without short-term loans and institutions 5 through 8 used DDA with short-term loans. Some institutions weeded the discovery pool and some did not; some discovery pools were automated via the YBP approval plan and some were entirely selector mediated. With so many variances in program durations, DDA business models, and collection strategies, it was a challenge to execute comparative analysis.

A broad comparative overview of each of the programs led to interesting discoveries. For example, one surprising discovery was the maximum expense for a title. Most institutions had a price cap of around $250, but it became apparent that the cap did not hold true. Some institutions spent far more than the authorized cap. When providers were questioned about this, it was explained that an e-book may come into the discovery pool under the cap but the publisher might change the price between the time it entered the discovery pool and the time it was triggered. The new Ebook Central platform has a mechanism to avoid this problem in the future.

An important measure of a successful DDA program is whether e-books continue to receive usage after the initial trigger for purchase or loan, that is, to determine if a patron's usage-initiated trigger is random act or an indicator of need that leads to sustained use. To do that, the following methodology was employed. First, data of the first complete year e-book cohort from the trigger report was identified. Then we tracked their usage over the program duration as reflected in the ProQuest "All Titles" reports delineated by program year. Depending on purchase license and institution's DDA model, a triggered title in the cohort may experience: (1) just STL(s), (2) purchase under a particular license, or (3) STL(s) plus purchase.

During the process, some issues were encountered. First, it was difficult to accurately evaluate use over time for programs that weed their discovery pool for the reason that if a title is not purchased, it could be weeded despite having gone through STL(s). Such titles were not given the time to perform. Another issue was that some DDA programs have a selector-mediated approach rather than an entirely automated DDA selection for the discovery pool. As such, the initial purchase decision was not based on patron use to begin with.

In this study several measures were applied in order to examine e-book use over time. First was the percentage of e-books that continue receiving use over program duration. All e-books in their respective program's first-year cohort received use in their trigger year. It was found that, overall, there was a decline in use of the titles after the trigger year; however, over time, close to 20% of e-books received steady usage, even in their fifth year. Another measure to gauge use over time was the number of user sessions, which demonstrate the user demand. It was found that the first-year cohort received use throughout the program duration, although it might vary by program size, and there was a decline in user demand over time. The third measure to gauge use over time was the user sessions/titles used, which shows the intensity of use of those used titles. The data shows that the intensity of use remained strong, with at least three user sessions per title even into the fifth year.

One consideration when implementing a DDA program is whether or not to include STLs. The Kent State program initially adopted the direct -urchase DDA model without STLs. However, a question remains about whether it would make sense to incorporate STLs for the program. As shown in the program survey of the eight institutions, the DDA programs vary to a great extent due to many variables. In order to perform a controlled study, an apple-to-apple comparison of DDA programs is necessary. This could be accomplished based on actual e-book use under different "what-if" scenarios to indicate the best option for each program.

Four institutions' DDA programs were chosen for the comparison. Those included were two straight DDA programs and two that incorporated STL programs. The choices were based on the following considerations/factors in order for the programs to

be comparable: (1) the programs were not weeding their DDA pool, (2) the programs have comparable cohort and sample size, and (3) the minimum program duration for the cohort had to be at least three years.

Three common scenarios were chosen for comparison: Scenario 1 (S1) , direct single-user purchase option (SUPO) purchase; Scenario 2 (S2), one 1-day STL followed by SUPO purchase; Scenario 3 (S3), three 1-day STLs followed by SUPO purchase. These scenarios were the most common purchase licenses among the eight study programs.

The formula logic for different scenarios was based on actual e-book use data. Step 1 was to determine which titles were STL eligible based on Available License data in the Trigger Report for each title. In Step 2 the Raw Trigger Total based on usage was calculated. The ProQuest ebrary trigger threshold is 10-10-1-1-1, which means any copy, print, or download may serve as a trigger, while 10 page turns or 10 consecutive minutes of use may also serve as a trigger. So for titles that were STL eligible, the Raw Trigger Total for each title can be calculated as follows: N of Views/10 + nCopy + nPrint + nChapter DL + nFull DL. Step 3 involved calculating the number of triggers based on Step 2 combined with User Sessions for each scenario. To do this the raw trigger total was compared with the corresponding number of user sessions to determine the likely number of triggers that could be applied to each of the different scenarios.

For example:

1 trigger = a purchase in S1; 1 STL in S2 and S3

2 triggers = a purchase in S1; 1 STL + purchase in S2; 2 STLs in S3

3 triggers = a purchase in S1; 1 STL + purchase in S2; 3 STLs in S3

4 or more triggers = a purchase in S1; 1 STL + purchase in S2; 3 STLs + purchase in S3

In Step 4 the cost of each title was determined based on the single-user publisher list price for non–STL eligible e-books, the mapped outcome of STL eligible e-books from Step 3, the publisher's STL rate and single-user publisher list price. Two examples follow of an e-book published by Princeton University Press, whose 1-day STL rate is 30% of list price and the list price is $100:

- If the e-book had 2 triggers, then
  - In S1, it would be purchased at list price at $100.
  - In S2, it would fall in 1 STL + purchase for a cost of $130.
  - In S3, it would be 2 STLs for a cost of $60.
- If an e-book had 4 or more triggers based on its usage, then
  - In S1, it would be purchased at list price at $100.
  - In S2, it would fall in 1 STL + purchase for a cost of $130.
  - In S3, it would be 3 STLs plus purchase for a cost of $190.

Finally, in Step 5 the total cost and percent of STL vs. purchase was calculated under each scenario for each program. The results were summarized in a side-by-side comparison of STLs vs. purchase. Results showed that for the S2 one 1-day STL followed by purchase, the majority of the initial cohort would have been purchased over the program duration. The ratio of the purchases ranges from 78% to 95% with varied program duration from three to five years. Under S3's three 1-day STLs followed by purchase, results showed that for three programs, over two-thirds of the initial cohort would have been purchased over the program duration and over half of the initial cohort would have been purchased for the fourth program with a three-year program duration.

In a more detailed breakdown of STL vs. purchase for the programs, results show that for institution 1 AND institution 2:

- Under 10% of their initial cohort would still be in 1 STL stage.
- Over 20% of their initial cohort were NOT STL-eligible per publisher license.
- Over 90% of their initial cohort would have been purchased under S2.
- About 70% of their initial cohort would have been purchased under S3.
- Any additional triggers of the cohort titles that were in the STL stage may lead to purchase down the road with additional use.

Results show that for institution 8 over 50% of e-books in the initial cohort would have been purchased after going through 3 STLs, 84% of the cohort would have been purchased under S2, and 68% under S3. Institution 5 was an outlier with the highest portion of 23% of their initial cohort in 1 STL stage among all programs.

Results of the cost comparisons of the three scenarios across programs showed that, with the exception of institution 5, S1 (direct SUPO purchase) is the most cost-effective across the board based on the actual use of the initial cohort over the program duration. For institution 5, even given the partial third-year use data among three scenarios is minor with the consideration that S1 would not incur any further cost with further use of the titles.

So what did we learn from this exercise? First, we learned that there is value in the longitudinal study. The element of time shows the bigger picture of e-book use. We learned that once an e-book is used, it tends to incur continued use in subsequent years, which has consequences for the overall cost of the short-term model and can help formulate weeding protocols. We also learned that the analysis process is data-rich, very time-consuming, and labor-intensive. And we learned that determining the best value and best practice is difficult to assess due to the fact that each DDA program varies and each library has its own collection strategies. For example, some prefer ownership vs. lease or a combination of both, some libraries weed the discovery pool, and some have not.

When we look at the DDA programs in terms of the scenario analysis, we know that the short-term business model may make sense initially, but for three out of the four institutions the straight-purchase DDA model had a greater return on investment for owned content.

Demand-driven acquisitions represents a shift away from the traditional "just in case" collection development model. Based on program variability, we would argue that determining which DDA model is best for a library really depends on that library's philosophy, mission, and collection strategy. Today, more than ever, libraries are becoming a "just in time" service point that meets the immediate needs of the user. Nonetheless, for many libraries traditional collection concepts continue in tandem, balancing ownership vs. lease. Examining actual evidence-based use data and employing longitudinal formulaic analysis will provide a means for assessing whether or not the program meets library goals and will help detect patterns that predict cost and return on investment.