Tightening the Core: Using Circulation and Cost History to Reduce Spending on a Research Library's Central Approval Plan

Richard Entlich
Cornell University, rge1@cornell.edu

Maureen Morris
Cornell University, mm342@cornell.edu

Follow this and additional works at: <https://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>.

Richard Entlich and Maureen Morris, "Tightening the Core: Using Circulation and Cost History to Reduce Spending on a Research Library's Central Approval Plan" (2009). *Proceedings of the Charleston Library Conference*.
<http://dx.doi.org/10.5703/1288284314735>

TIGHTENING THE CORE: USING CIRCULATION AND COST HISTORY TO REDUCE SPENDING ON A RESEARCH LIBRARY'S CENTRAL APPROVAL PLAN

Richard Entlich (rge1@cornell.edu) - Research and Assessment Librarian, Cornell University Library, Ithaca, NY

Maureen Morris (mm342@cornell.edu) - Reference Librarian, Cornell University Library, Ithaca, NY

ABSTRACT

Facing reductions to the materials budget due to the economic downturn, Cornell University Library has sought savings in nearly all areas of acquisitions. The central monographic approval plan was a significant concern, because it represented more than 7% of the total university appropriation for the library materials budget, and had been running at a deficit for years. The library's Collection Development and Research & Assessment Unit worked together to adjust the approval plan profile to meet budget targets. Research & Assessment combined multiple years of circulation data with cost data in a manner that allowed breakdown of average cost per item and cost per circulation, by publisher and by LC class. Collection Development staff devised methods to use the data to rapidly implement cuts in a way that would maintain traditional areas of collection strength, while minimizing the impact on selector workload and user needs. This paper explores the interdepartmental collaboration, the techniques that were developed to merge circulation data from the ILS with cost data from the approval plan vendor and to present the massive dataset in ways that were meaningful and easily understandable to selectors, as well as the ongoing process of monitoring the impact of the cuts.

History of the Central Approval Plan at Cornell University Library

Cornell University Library established a central approval plan with Yankee Book Peddler in 1991. Prior to this Cornell had a few small approval plans including Oxbridge and some small science plans, but the majority of material was purchased as firm orders or monographic series standing orders by selectors managing individual funds.

1980s	-90% firm orders or monographic standing orders -10% on approval (Oxbridge Plan and science plans)
1990s	-60% firm orders and standing orders -40% approval
2009	-45% firm orders (standing orders eliminated) -55% on approval

Purchasing distribution between firm orders/approval plan, 1980s-2009.

An approval plan of university press and selected trade publishers was set up in 1991. For the first four years the approval plan was funded from the budget lines of the individual selectors. In 1995 a central fund was created in which the individual selector funds were taxed based on the percentage of the approval plan they had funded in the previous years. The approval plan remained relatively stable throughout the past 20 years although a few publishers were added and dropped. No major review of the plan took place during this time.

Almost from the beginning the central approval plan budget line was overspent, with the over expenditure growing through the years. This was potentially attributable to a number of factors but the most likely cause was that the original allocation was simply not adequate to support the plan. This issue, coupled with recent economic pressure, led to a need to re-examine the plan. In 2009 the library materials budget sustained a 5.3% reduction and endowment payouts were reduced by 15%. The initial proposal for bringing approval plan expenses under control was to cut spending by 30%. This amount would take into account the over expenditure, the current budget cuts, and the possibility of further materials budget cuts next year.

There was an overriding question that surfaced during this process. Why target the approval plan? Since the approval plan presumably contains the core set of heavily used material, why not look at the collection as a whole and possibly protect the approval plan? While everyone agreed it would be ideal to examine the whole collection this way, time was an overriding factor. Approval plan spending had been out of balance for some time, and a comprehensive review of collection spending could not be completed in time to meet budget deadlines.

Fortunately these decisions did not have to be made without data to support them. We were able to turn to the library's recently-formed Research & Assessment Unit for help with data gathering and analysis.

Background on the Cornell University Library Research & Assessment Unit (RAU)

Cornell University Library formed its Research & Assessment Unit (RAU) in 2007 as a concrete manifestation of its commitment to the practice of evidence-based decision-making. Previously, the library had a research unit, but its focus had been on digital imaging and digital preservation, during the time when these were more experimental activities than they are today.

The old research unit's former responsibilities had also encompassed data collection in support of internal and external reporting. This included producing the library's annual statistical report, and meeting the data needs of the university's central administration, library associations (e.g., ARL), federal agencies such as NCES (National Center for Education Statistics), and private publishers like US News & World Report, Barron's, etc. These activities remained with RAU.

No additional staff were hired. Existing staff had good research and technical skills, but little familiarity with formal assessment techniques. Much of the unit's first year was spent learning the fundamentals of evaluation and assessment, partially through self-study and training, and partially by the experience of conducting a variety of assessment projects. During its formative period, RAU staff learned to use new tools and methods, including surveys, focus groups, interviews, observation, statistical analysis, text analysis, and data manipulation. The unit worked to improve the sophistication of its efforts, and to establish credibility within the library.

The unit came together just in time to help the library face the economic downturn of 2008. As the economy worsened, more of the library's priorities centered on response to fiscal pressures, including staff reorganization, closing and consolidating unit libraries, and dealing with materials budget constraints. All these challenges have been more effectively addressed using a rational approach undergirded by solid evidence.

Collaboration between Collection Development and RAU

Collection Development (CD) approached RAU early in 2009 for assistance in gathering data that would facilitate making a significant cut to the central monographic approval plan. A group

of six selectors from different subject areas met with two RAU staff to develop a work plan and timeline, and to define the project's deliverables.

CD wanted to complete the process of making as much as a 30% reduction in spending within six months. A key reason to involve RAU was the desire to factor both cost and usage considerations into reworking of the approval plan profile. This greatly complicated the data gathering and analysis tasks, because no single source of cost and usage data for approval plan items existed.

Eventually, the request to RAU took the following form

- assemble cost and usage data for central approval plan items
- merge cost and usage data to determine cost per use, by publisher and LC class
- examine the impact of changing the plan's price cap¹

Data challenges

Challenge #1: Gathering and manipulating the data

RAU's initial challenge was to gather the requested data. Circulation data going back about 10 years is readily available from Cornell's Voyager ILS, but cost data for individual central approval plan items is only available in reports from the approval plan vendor. We began with approval plan data and tried to match it with records for the same items in the ILS. However, Yankee's online customer data system, GOBI (Global Online Bibliographic Information), had some limitations. Its reporting module could extract summary data going back about four years, but not at the title level, which we needed to match records in the ILS. Its search module could generate title level data, but no more than 1000 records per query, which is less than one month's data for the library. Furthermore, GOBI's records all used ISBN-13s, while many older Cornell records had only ISBN-10s.

We then tried to identify relevant approval plan records in the ILS itself. Not only do we have more control over the ILS, but ILS data had more historical depth than the Yankee data, going back to the 1990s. Unfortunately, this too was unsuccessful, as we discovered that until fairly recently, it was not uncommon for catalog records accompanying approval plan items to be overlaid with more complete records from other sources. In the process, data identifying these records as having originated with the approval plan was lost.

Ultimately, we turned to Yankee for help, and they agreed to provide the data we needed in the form of custom reports. We requested, and received title level data on shipped items (returns excluded) including title, author, publisher, ISBN, LC class, ship date, invoice date, list price, and net price. Yankee supplied the data in five separate spreadsheet files (year by year, from 2005-2009), including ISBN-13s and ISBN-10s for each item.

Data for the five years was combined into a single Excel spreadsheet, and consisted of over 60,000 records. The ISBN data in this file was then used to extract usage data from the ILS, a surprisingly smooth process, marred only by some ISBN duplication by format (print and eBook with the same ISBN), and in a few cases, reuse of ISBNs for entirely different titles. The last step was to aggregate the 60,000 plus records into a much smaller number of publisher and LC

¹ The price cap analysis was conducted, but ultimately wasn't a factor in decision making.

class categories, since those were to be the primary bases for adjusting the approval plan profile to reduce overall expenditures.

Microsoft Excel provides a number of different ways to consolidate data, but the most flexible and powerful is through the use of pivot tables. In a pivot table, each column represents a data field which is used to summarize multiple rows of data with common elements. A major obstacle to effective use of pivot tables with our data was the considerable amount of “noise” in the Yankee data’s publisher name field. Pivot tables will only consolidate rows with identical contents.

Unless corrected, variations caused by misspellings, abbreviations, punctuation differences, and name changes would have prevented complete consolidation by publisher and blunted the usefulness of the summary. This was the most tedious part of the data manipulation, because there was no way to automate the necessary corrections. Ultimately, nearly 200 variant publisher names (out of about 550 in the initial data) were filtered out or combined with others, leaving a more manageable group of 350 for final analysis.

Challenge #2: Presenting the data

The final challenge for RAU was to present the combined cost and usage data, consolidated by publisher and LC class, in as transparent a manner as possible. RAU created Excel workbooks with individual worksheets containing different “views” of the data. A portion of one view of the publisher data found to be particularly useful for decision making is shown in figure 1.² Among other things, it contains the following data fields:

Column heading	Data description	Data type	Data source(s)
A	Publisher name, post cleanup	Raw	Vendor
B	Total number of items from this publisher	Raw	Vendor
C	Total circulations for this publisher	Raw	ILS
D	Total cost of items from this publisher	Calculated	Vendor
E	Cumulative cost of all items, up to this level of circulation	Calculated	Vendor
F	Cumulative percentage of total cost, up to this level of circulation	Calculated	Vendor
G	Average cost/item	Calculated	Vendor
H	Average cost/circulation	Calculated	ILS and Vendor
I	Average circulations/item	Calculated	ILS

² Spreadsheets organized around LC classes had an identical structure, with column A representing LC class instead of publisher name.

	A	B	C	D	E	F	G	H	I
1		Tot	Tot			Cume	Avg	Avg	Avg
2		items	hist	Tot	Cume	%	cost/	cost/	circs/
3	Publisher	purch	chgs	cost	cost	cost	item	use	item
148	HOOVER INSTITUTION	25	15	\$351.35	\$762,917.02	28.41	\$14.05	\$23.42	0.60
149	UNIV OF BRIT COLUM	177	110	\$12,894.38	\$775,811.40	28.89	\$72.85	\$117.22	0.62
150	LOYOLA PRESS	8	5	\$127.06	\$775,938.46	28.90	\$15.88	\$25.41	0.63
151	STEERFORTH PRESS	16	10	\$312.39	\$776,250.85	28.91	\$19.52	\$31.24	0.63
152	UNIV PR OF COLORA	62	39	\$2,256.26	\$778,507.11	28.99	\$36.39	\$57.85	0.63
153	NORTHWESTERN UN	123	78	\$3,691.22	\$782,198.33	29.13	\$30.01	\$47.32	0.63
154	QUANTUCK LANE PR	11	7	\$411.94	\$782,610.27	29.15	\$37.45	\$58.85	0.64
155	CAMDEN HOUSE	58	37	\$3,585.32	\$786,195.59	29.28	\$61.82	\$96.90	0.64
156	UNIVERSITY PR OF F	282	183	\$11,916.94	\$798,112.53	29.72	\$42.26	\$65.12	0.65
157	UNIV OF WISCONSIN	191	124	\$6,567.64	\$804,680.17	29.97	\$34.39	\$52.96	0.65
158	COPPER CANYON	52	34	\$809.93	\$805,490.10	30.00	\$15.58	\$23.82	0.65
159	TRANSACTION	204	134	\$8,346.80	\$813,836.90	30.31	\$40.92	\$62.29	0.66
160	BOA EDITIONS	3	2	\$47.02	\$813,883.92	30.31	\$15.67	\$23.51	0.67
161	HOLMES & MEIER	3	2	\$87.95	\$813,971.87	30.31	\$29.32	\$43.98	0.67
162	OREGON STATE UNIV	27	18	\$583.38	\$814,555.25	30.34	\$21.61	\$32.41	0.67
163	POPULAR PRESS	6	4	\$269.71	\$814,824.96	30.35	\$44.95	\$67.43	0.67
164	UNIV PR OF KENTUC	159	106	\$5,097.15	\$819,922.11	30.54	\$32.06	\$48.09	0.67
165	ELSEVIER	366	247	\$42,061.21	\$861,983.32	32.10	\$114.92	\$170.29	0.67

Figure 1

Figure 1 shows a small portion of the worksheet. Because the data is sorted in ascending order by average usage (average circulations per item, shown in column I), the publishers with the lowest average circulation appear first. By also showing the cumulative total cost (column E) and cumulative percentage cost (column F), selectors could easily choose a target dollar amount or percentage of savings, and see which publishers fell into the low usage category represented by that amount of expense reduction.

For example, to achieve a 30% cut solely from the lowest usage publishers, all the publishers from row 158 and up would have to be removed from the plan.

Some selectors requested a graphical representation of this data. Initially, we plotted average circulation against the cumulative percentage of cost, as shown in figure 2.

Figure 2

This graph reveals general trends. In order to simplify the presentation, the average circulation figures were rounded off to two decimal places, meaning that data points do not necessarily correspond to individual publishers. An interesting features of this graph is the large, nearly vertical displacements, which represent significant costs associated with just a few publishers.

To improve the utility of this graph, we created figure 3, a higher resolution version of the same data, limited to the bottom 30% of cumulative cost, representing only the lowest usage publishers. The relatively low-use publishers corresponding to significant percentages of total approval plan spending were labeled for easier identification. Subsequently, RAU was asked to generate further breakdowns of some of these publishers by LC class to provide a more granular look at usage for these large publishers.

Figure 3

After delivering the final data, along with documentation on its use and interpretation, and caveats about its limitations, RAU signed off on its role in the project, leaving CD to apply considerations of collection building priorities and selector workload to the final decision-making process.

When data meets reality

The approval plan reduction was not happening in isolation. There were a number of other factors affecting the budget and selectors' time including:

- Individual selector budget lines were reduced an average of 8%.
- A journal cancellation project was taking place.
- Standing orders were being eliminated.
- Several key staff retired and selector duties were redistributed.

These factors contributed to the initial negative reaction from selectors to such a radical cut of the approval plan. A cut of 30% would have eliminated 155 publishers, many of them prominent publishers in certain disciplines. In addition to the potential loss of some core content, the main concern of selectors was the time it was going to take to keep track of these publishers individually. There was also discussion about whether to privilege the university press publishers, which had been the core of the original approval plan.

Meetings of all selectors and smaller subject groups were held and a wiki was employed to gather as many of the concerns as possible. Ultimately a revised proposal was put forth to concentrate on the most expensive, highest output publishers that still fell within the lower end of the circulation statistics. This resulted in a much shorter list of 20 presses to move to "slip-

only” status which still allowed us to realize a 21% reduction in central approval plan spending. The decision was made to retain all of the university presses though this was not unanimously supported. This revised proposal was accepted and implemented in July 2009.

Implications

While it is still very early in the process, we have already begun to see the consequences of the changes. As hoped, some lower use material has been identified and is no longer being selected for our collections. However, some items that previously came through on the approval plan have been identified by selectors as important to collect and these now have to be funded in other ways. In some instances individual selectors are making the choice to purchase items on their subject funds. This in turn means that some firm order items or journal subscriptions will be dropped from subject lines in order to free up the necessary funds.

So even though the approval plan was the focus of the initial cuts, it is the entire collection that is being affected. Selectors are balancing the desire to keep some of the previous approval plan material against the purchases they normally would have made from their subject lines. The time it takes to make these decisions, as well as the increase in slips to sort through and publishers to track, has led to some increase in individual selector workload. Time is needed to determine if this is a sustainable burden and whether further adjustments need to be made to the plan.

Next Steps

In addition to monitoring the effects of the decisions we have made in the previous year, we are also interested in taking a broader look at our collection use. A new task force is working to analyze circulation statistics for the entire print monograph collection. Circulation is not the only basis for making collection decisions and as such the group undertaking this project will consider usage in the context of the various disciplines and their representation at Cornell, taking factors such as language and potential readerships into account. As stated in the task force’s charge, the ultimate goal of this process is to “propose an intellectual framework for thinking about collection ‘use’ in the 21st-century research library, synthesizing the group’s analysis of the [library’s] usage data, and providing recommendations as to how usage should inform selection and allocation decisions.”

References

- Brush, D. (2007). Circulation analysis of an engineering monograph approval plan. *Collection Building*, 26(2), 59-62.
- Case, B. D. (1996). Approval plan evaluation studies: A selected annotated bibliography, 1969-1996. *Against the Grain*, 8, 18-21+.
- Glazer, Gwen. (2009) Library scales back on books, journals, databases. *Cornell Chronicle Online*, April 14, 2009, <http://www.news.cornell.edu/stories/April09/LibraryBookCuts.gg.html>
- Kingsley, M. S. (1996). Circulation statistics for measuring approval plan effectiveness. *Against the Grain*, 8, 1+.
- Lorenzen, E., & Kapusta, S. (2007). Long live the approval plan, even in a time of financial constraints... *Against the Grain*, 19(3), 40, 42.
- Tucker, J. C. (2009). Collection assessment of monograph purchases at the University of Nevada, Las Vegas libraries. *Collection Management*, 34(3), 157-181.