

# Buying by the bucketful: a comparative study of e-book acquisition strategies

Revised and expanded article updating on a presentation originally given at the 34th UKSG Conference, Harrogate, April 2011

The University of Liverpool (UoL) has been subscribing to and purchasing e-book packages since 2006. This article describes UoL's experiences of evaluating the use and value of such packages, and considers whether alternative purchasing strategies – single-title selection (also known as 'firm orders'), patron-driven acquisition (PDA) or evidence-based selection (EBS) – might offer a more cost-effective way to meet users' needs. The discussion is informed by evidence from the University of Liverpool COUNTER usage statistics and includes modelling of the costs that might result from observed usage patterns if a PDA purchasing model were applied. Criteria for evaluating the success of e-book packages are discussed, including a consideration of how value for money can be calculated for different e-book purchasing models. An outline is provided of the implications for managing acquisitions budgets when e-book collections form part of the acquisitions strategy.

## Introduction and context

The University of Liverpool (UoL) was founded 1881. It is a research-intensive institution and is a member of the Russell Group. The University carries out teaching and research in most disciplines: medicine, biological sciences, veterinary science, physical sciences, engineering, social sciences and law, management and the humanities. It offers a number of online degree programmes, with the greatest numbers in management and IT programmes.

Having previously been poorly funded relative to its peer institutions, the UoL library underwent a period of budgetary growth from 2006 to 2009. Starting from a modest subscriptions base meant that journal 'big deals' were an attractive proposition. The library has had much experience of studying the usage reports from its e-journal collections. It has concluded that packages nearly always give greater value for money than single-title subscriptions, and that the journals that had historically been selected for subscription were often not the most well-used<sup>1</sup>.

Book acquisition had historically been very selective at the UoL library. Books were only purchased when specifically requested by academic staff and, to a lesser extent, by students. Although the library has subject librarians, they did not speculatively build collections, neither did the library participate in the sort of approval plans with book vendors that are (or were) so common in the US.

The library began to be interested in e-books in 2005. An early experiment with a handful of titles had little impact, but in 2006, the library began a subscription to ebrary's Academic Complete subscription bundle. Usage was immediately impressive and it quickly became established as one of the library's best value subscriptions. It also seemed to precipitate a sea-change in user expectations. Previously, users had not demanded e-books, but as soon as a critical mass of e-books became available (the collection contained about 30,000 titles at the time and over 70,000 now), e-books were regularly appearing in catalogue



TERRY BUCKNELL  
Electronic Resources  
Manager  
Sydney Jones Library  
University of  
Liverpool

52 search results, and “we need more e-books” became a regular demand. The library started to purchase single e-books in response to these demands but these were on aggregator platforms and came with digital rights management (DRM) that restricted how many users could access a title at once, and/or how much of the book could be copied, printed or downloaded for offline reading. This quickly became a source of annoyance for users, and a source of confusion for the library staff that had to support them.

As the UoL library’s funding improved, the procedures for selecting books did not change substantially. As a result, the library regularly found itself with end-of-financial year surpluses that needed to be spent quickly on ‘goods’ that it knew would be ‘received’ promptly. Purchases of electronic resources collections were an obvious way to spend to these funds. Journal back-files were the traditional end-of-year electronic purchase, but COUNTER Report JR1a (and later JR5) showed that these were not actually very well used. Rather, their benefit was in allowing the library to clear its shelves of previously purchased print runs, using managed disposal through the UK Research Reserve ([www.ukrr.ac.uk](http://www.ukrr.ac.uk)). Instead, e-book collections were acquired on the assumption that these would maximize the benefit to students and staff. Because these collections were generally on publisher platforms without DRM, the library felt sure that they would be popular with users as long as the collections were relevant to the needs of users.

## Evaluating usage and impact of e-book collections

As budgets reduced again, it was imperative to examine the usage of these packages to ensure that they were giving the best value, and to prioritize which packages should be purchased on an annual basis. The library had been purchasing Springer’s complete collections of e-books since July 2008, and the first purchase included all titles published between 2005 and 2007 too. In 2010, the library engaged on a research project with Springer to examine user attitudes to e-books through an online survey<sup>2</sup>, and to examine usage patterns through analysis of COUNTER usage reports<sup>3</sup>. The patterns of usage that had been previously reported have continued. Total usage of full-text content on the platform has continued to grow (with a seasonal variation, of course), and several recent months have seen more e-book chapters downloaded than journal articles (Figure 1).

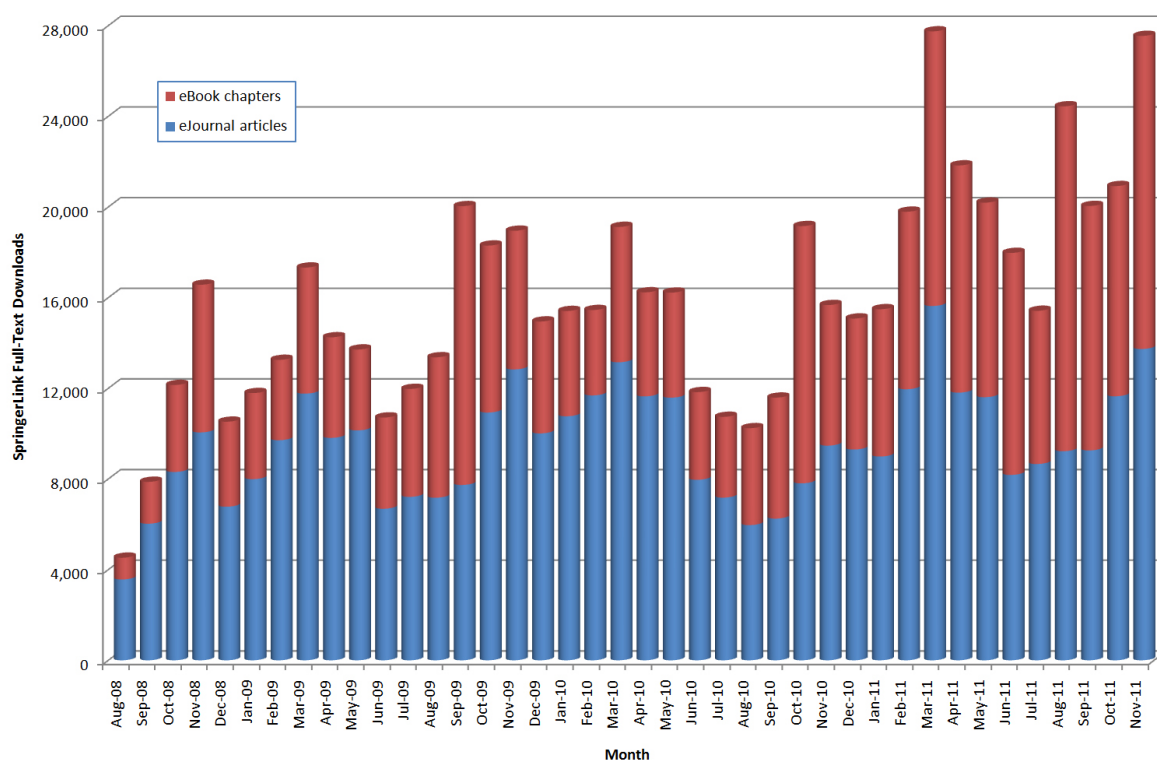


Figure 1. UoL’s full-text download usage statistics from the SpringerLink platform between August 2008 and November 2011, showing the growth in total usage and the increasing proportion of book chapter downloads.

53 Once UoL launched its EBSCO Discovery Service in October 2010, the usage of SpringerLink journal articles declined a little at first compared to 12 months previously. This may have been due to users being unfamiliar with the new system, or it may have been because users were more likely to follow a PDF link to the article in a subscribed EBSCO full-text database where available, rather than follow the library's SFX link to the article on SpringerLink. But it might have been that users who would previously have read journal articles were now reading e-book content instead. Since February 2011, journal usage has recovered and is now growing. Use of e-books has grown throughout, and has accelerated in recent months, apart from in October 2011 in comparison to October 2010, which had unusually high usage (Figure 2).

It is tempting to think that the discovery service is helping students to find content that is more suited to them. Previously, students probably tended to search the library catalogue for known items, and when they needed to find content on a topic, they were encouraged to search databases and will mainly have retrieved e-journal articles. Now, they search the discovery service for topic-based searches and they see e-books and e-journal articles together in their results set. Perhaps they prefer to read e-books because the content in them is more appropriate for their level of study.

At present, the UoL library uses basic, Springer-supplied MARC records which contain book-level metadata only. The library will be replacing these with enhanced MARC records (via OCLC) that contain tables of contents, so it will be interesting to see if this more granular metadata generates an even greater level of usage.

For e-journals, librarians are accustomed to calculating the cost per download by dividing a single year's payment by a single year's usage. A different approach is needed to compare the value for money for an e-book package subscription against that of a purchased e-book collection and against that of single e-books purchased on a platform: calculate the accumulated expenditure (since the start of the subscription or purchases) divided by the accumulated usage.

For subscriptions and annual collection purchases, the cost per download leaps each time a payment is made and then slowly returns to its previous level just before the next payment is made (Figure 3). At each subsequent payment, the increase in cost per download is

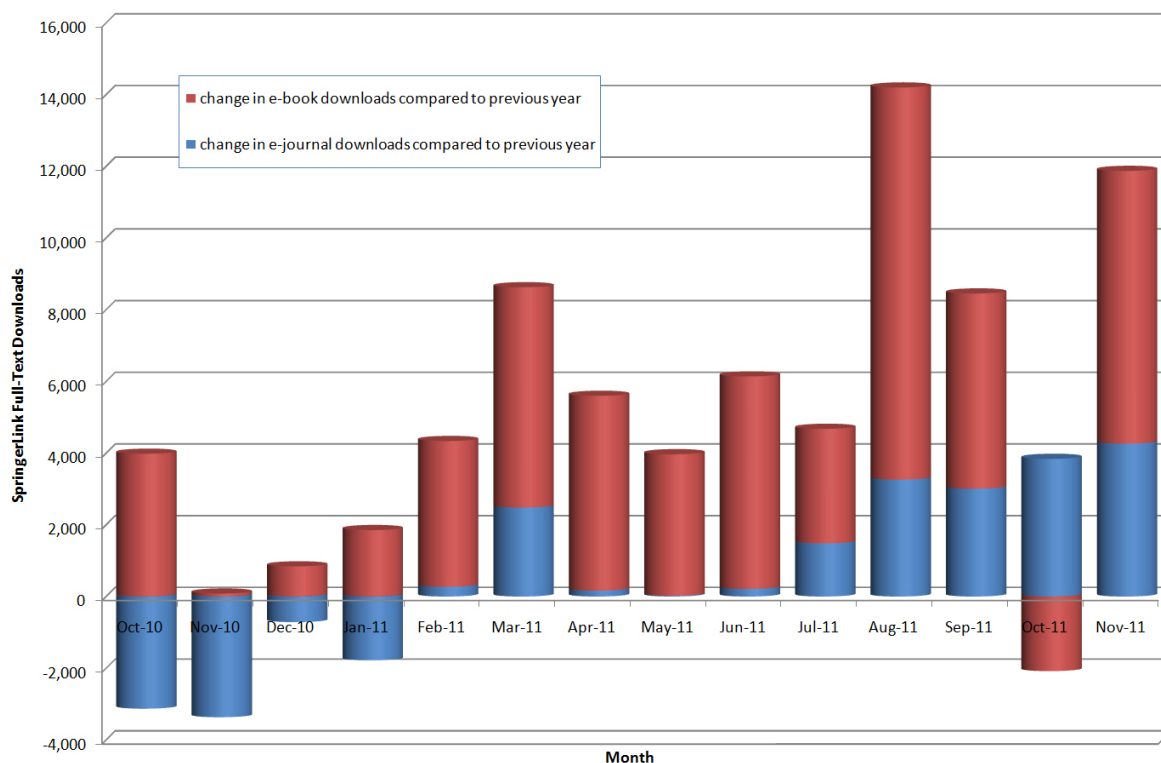


Figure 2. The difference between UoL's SpringerLink full-text journal article and book chapter downloads between October 2010 and November 2011 and those between October 2009 and November 2011. EBSCO Discovery Service was launched at the university in October 2010.

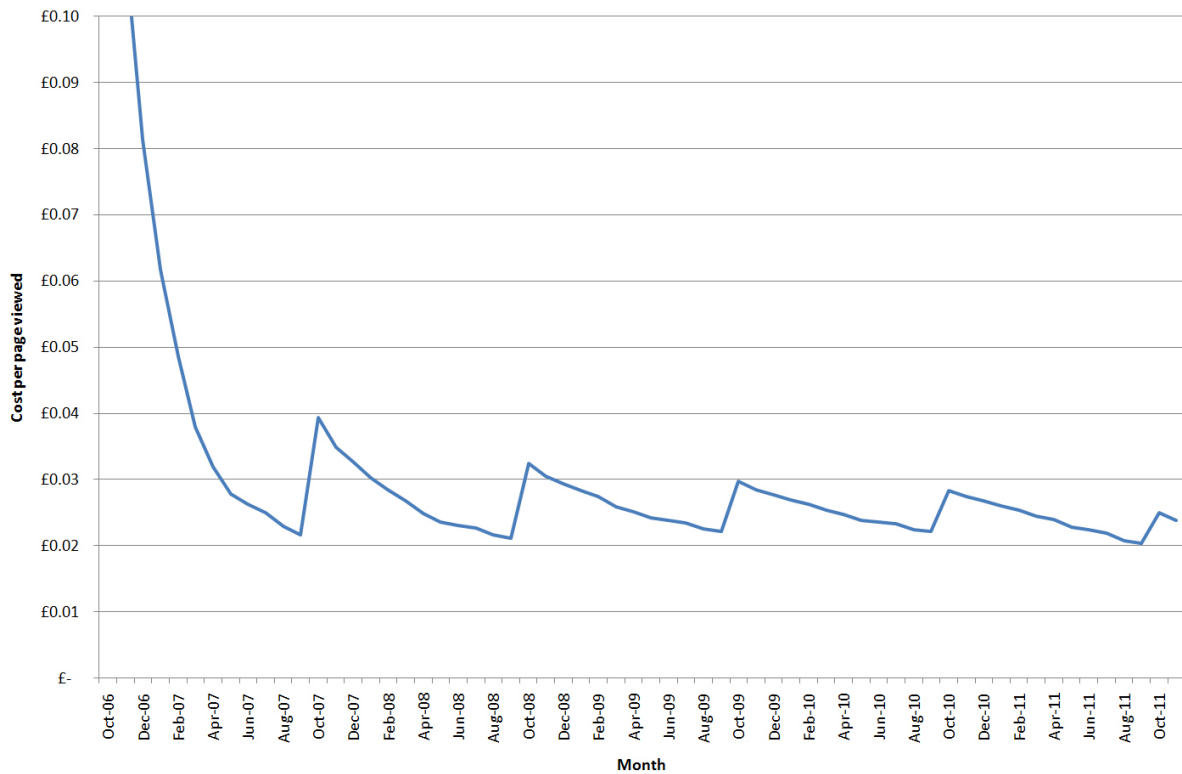


Figure 3. Cost per page viewed for an e-book collection that UoL started subscribing to in October 2006. The cost per page viewed varies between about 2p and 3p depending upon the position of the month in the subscription cycle.

relatively smaller because each payment becomes a successfully smaller proportion of the total payment made over all time.

For a single one-off purchase, the cost per download can only ever decline and will tend towards an asymptotic value. If collections are purchased annually then in combination they will show the same pattern as a subscription (Figure 4).

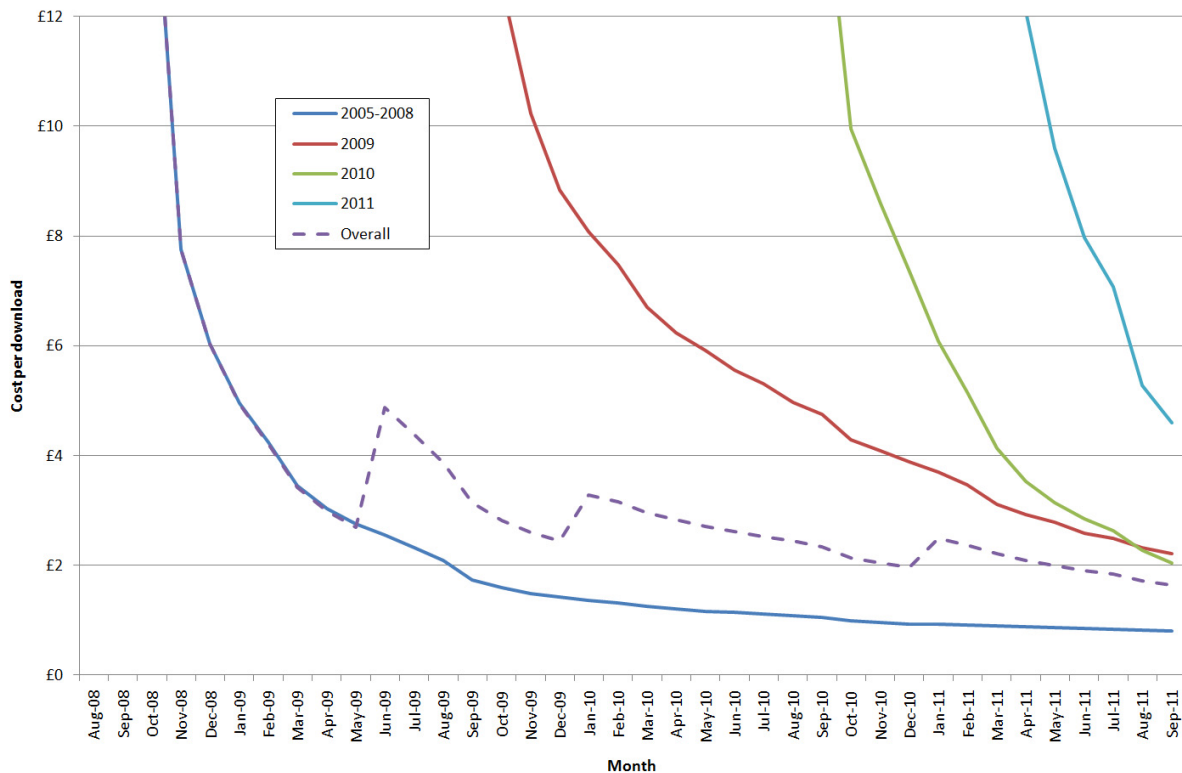


Figure 4. Cost per chapter downloaded for UoL's Springer purchases: the 2005-2008 collections purchased in July 2008; the 2009 collection purchased in July 2009; the 2010 collection purchased in January in 2010; and the 2011 collection purchased in January 2011. The earliest-purchased collection shows a cost per chapter of about 80p after three years. The dotted line shows the cost per chapter downloaded for all the purchased collections in combination. Since the most recent purchase it varies between about £1.60 and £2.40 as for a subscription (see Figure 3).

55 For a collection of single e-books purchased regularly through the year, the pattern is more complicated and will depend on how much is spent each month and when usage begins to occur for the titles purchased, but cost per use will still tend downwards towards a consistent value because the latest month's additional expenditure and additional usage will become a small proportion of the total (Figure 5).

Cost per download is of course not the full story, not least because it is so hard to compare cost per use on e-book platforms that function differently: Figures 3, 4 and 5 show cost per page viewed, cost per chapter downloaded and cost per book downloaded. Furthermore, a user can download a book chapter once from a DRM-free platform and use it multiple times over an extended period, but might have to download the chapter each time they wanted to use it on a DRM-enabled platform that only permits time-limited downloads.

The library might be able to spend its budget more advantageously if it could purchase the titles that would be well used for less than the price of the collection. To gauge the value of a particular package, it is therefore instructive to calculate the percentage of the collection that has ever been used, and, of course, this proportion will grow over time. For Springer collections the library has typically found that about 45% of titles will be used in the first year of ownership, about 65% in the first two years and about 75% in the first three years (Figure 6). The University of Liverpool has found that each year's Springer collection has behaved similarly, which gives confidence that if a one year's collection is used well, then so will future years'.

The library has plotted similar graphs for subject collections, and here differences do become apparent (Figure 7). It is assumed that these are a reflection of the institutional subject strengths and need rather than the quality of the collection in different subject areas. For the UoL, the best-used subjects correspond to those with extensive online degree programmes, large numbers of students on off-campus placements, or in subjects with high research rankings. Less extensively used collections correspond to subjects with lower research rankings or lower FTE counts, and with lower usage for their e-journal collections too. We tend to find that the same subjects perform best in other collections as well, and dominate the use of single-title e-book purchases. Consequently, the library has used this technique, and counts of how frequently titles are used, to decide which e-book collections (across multiple publishers and multiple subjects) to prioritize.

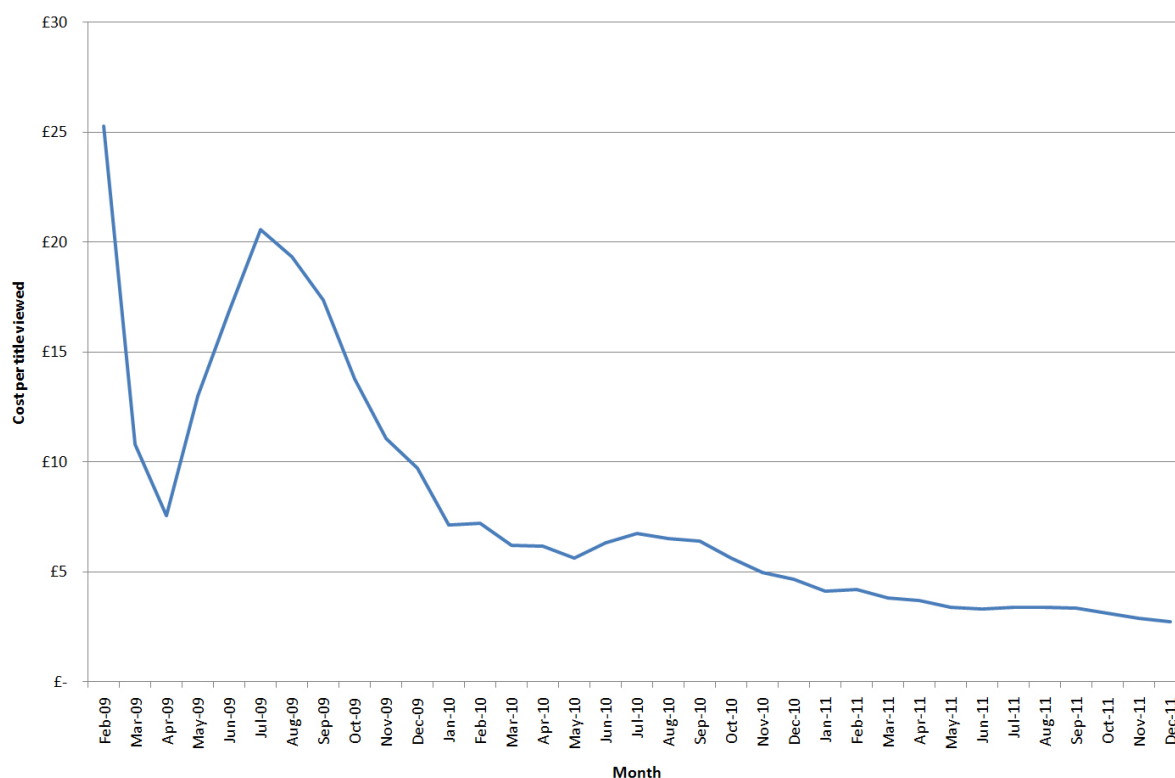


Figure 5. Cost per title viewed for an aggregator platform on which UoL has been purchasing single e-books every month since February 2009. These usage figures are from the platform's COUNTER BR1 report. The platform also started to offer a BR2 report in September 2010 but these figures are not included in the graph.

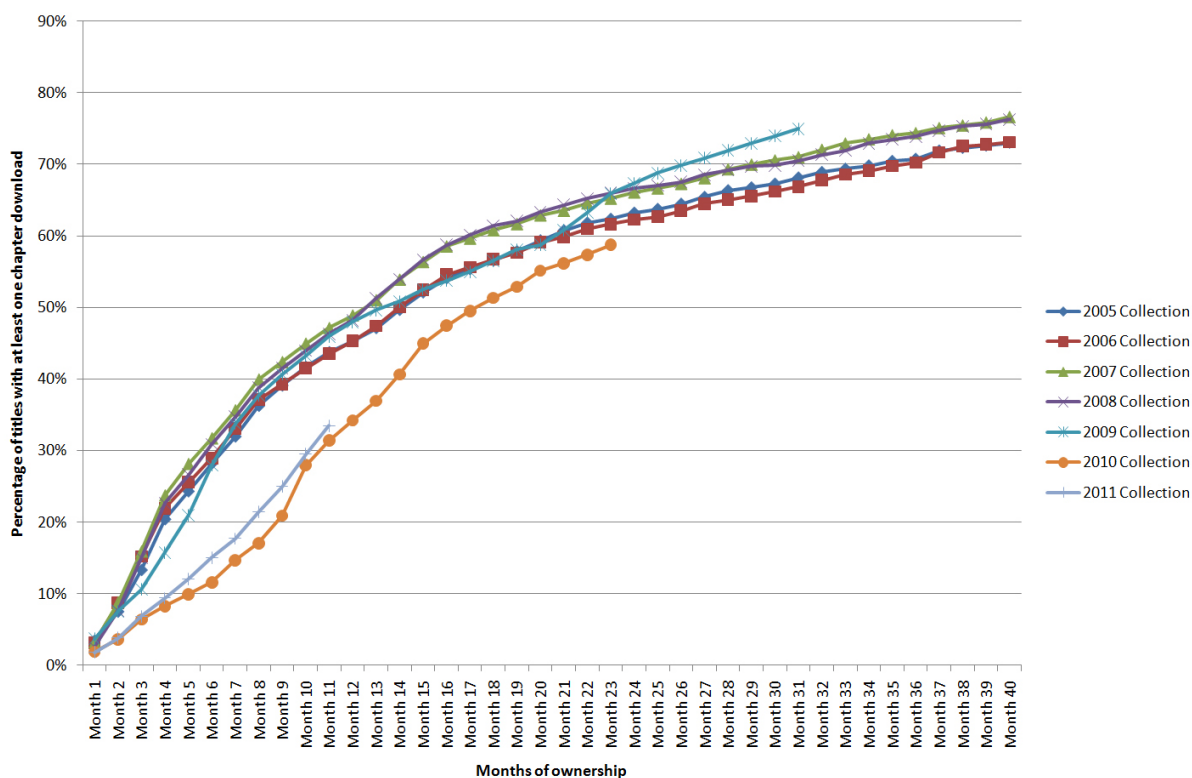


Figure 6. The proportion of titles in each of UoL's Springer collections that have seen at least one full-text download since the collection was purchased. Note that in 2010 the library started buying Springer collections at the start of the calendar year – when the collection had started to be made live – whereas in previous years collections had been purchased in July when much more of the collection was already online. This is reflected in the apparently slower take up of the 2010 and 2011 collections.

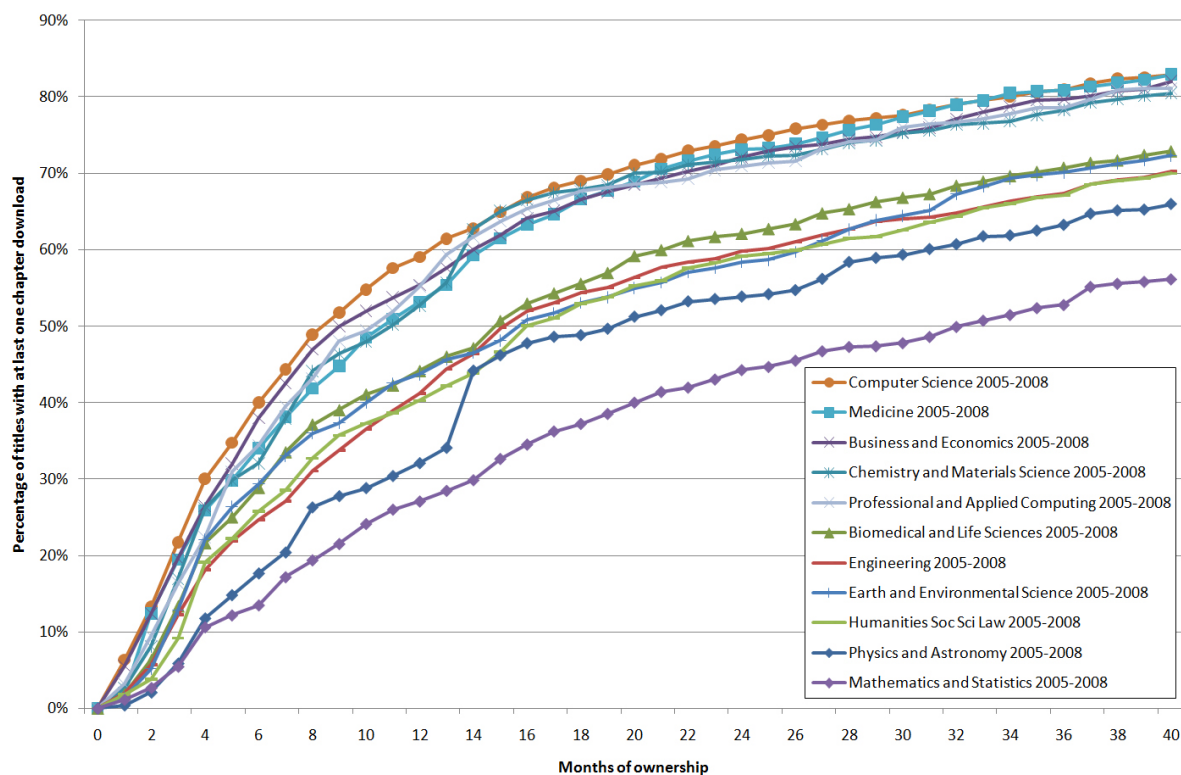


Figure 7. The proportion of titles in each of UoL's Springer 2005-2008 subject collections that have seen at least one full-text download since the collection was purchased. The collections of 2005-2008 titles were chosen as these had all been owned throughout the period under study.



## Patron-driven acquisition (PDA)

E-book packages can be criticized for perpetuating the practice of speculative collection-building, which is viewed as wasteful because it results in the purchase of content which is never used<sup>4</sup>. This was certainly true in the print world: every unused book bought through an approval plan wasted money that could otherwise have been spent on something better used, and took up valuable physical space in the library. E-books though are different: the only space that an e-book occupies is a catalogue record, and publishers can sell collections at a substantial discount off the total list price because they are not shipping physical items but merely activating access to files that already exist online.

Nevertheless, patron-driven acquisition is often touted as the most desirable way to acquire e-books because it ensures that the library only purchases e-books that are used, and indeed only those that are used above a certain threshold level. Some librarians worry that PDA removes the library's ability to build a balanced collection for future needs<sup>5</sup>. My concerns are rather more pragmatic (my library is not in the habit of speculative collection-building anyway):

- PDA tends to be synonymous with aggregator platforms at present. Aggregator platforms are synonymous with DRM and our users dislike DRM intensely
- evidence from the usage of e-book packages suggests that over a long period of time, a high proportion of collections will be used. If using a PDA model, purchases through PDA could cost more than purchasing packages that would serve those institutions' users just as well
- a major supposed benefit of PDA is that it allows library users to select from a huge range of possible titles. But experiences from other libraries are that PDA spends money very quickly. To keep expenditure down to manageable levels, a library has to restrict the size of the collection exposed to PDA. It does this by pre-selecting what titles it thinks users might be interested in, thereby denying access to other titles. So PDA still involves librarian selection after all
- conversely, if the library already provides access to a large collection of e-books, users may find themselves using PDA books that incur costs to the library when they could find that already-paid-for titles (that would incur no additional cost to the library) might meet their needs just as satisfactorily.

To test whether these fears were justified, UoL library modelled its usage reports from Springer and other packages as if a PDA model were being employed. Of course, real PDA models involve real-time monitoring of usage and they cause payment to be triggered according to well-defined event criteria. Conversely, COUNTER statistics only provide a monthly accumulation of data: ten chapters downloaded from one book in one month could involve between one and ten users on one to ten occasions.

At UoL, we decided to use the ebrary PDA model (which is very similar to EBSCO's model) as it does not use a 'rental' or 'short-term loan' option which would complicate the modelling. In ebrary's PDA model, a purchase is triggered if in a session:

- ten page turns in a title are made (not including front matter or index), or
- ten minutes are spent in a title in a user session, or
- one copy and paste is made, or
- one print is made.

For our modelling of Springer usage statistics, we made the assumption that any book with two or more chapter downloads would trigger a purchase. That is, if only one chapter was downloaded, we assumed that fewer than ten pages would have been read and no purchase would have followed. In reality, of course, that single chapter could have been downloaded,

printed or read in its entirety. Similarly, if more than one chapter was downloaded, the reality could have been that two users each downloaded one chapter and neither made enough use of it to generate a purchase under the ebrary model. To allow for the assumptions inherent in our model, we repeated the model with the threshold for purchase sequentially increased by one additional chapter up to a maximum of ten chapter downloads being required to trigger a purchase (Figure 8). The high proportion of collections used over time (Figure 6) would suggest that over time any PDA model is likely to exceed the collection price.

Even if ten chapter downloads would be required to reach the 'ten pages viewed' threshold that triggers a purchase under the ebrary PDA model, over a long enough timescale the PDA model would exceed the collection price. With a lower purchase threshold, the collection price would be exceeded sooner. As mentioned previously, other PDA models trigger in two stages: first a 'rental' or 'short-term loan' and then, after sufficient rentals, a full purchase would be triggered. No attempt was made to model this sort of PDA from our COUNTER reports but intuition suggests that although it would generate a lower level of purchases, these would be offset by additional, repeated rental charges (including rentals prior to eventual purchase).

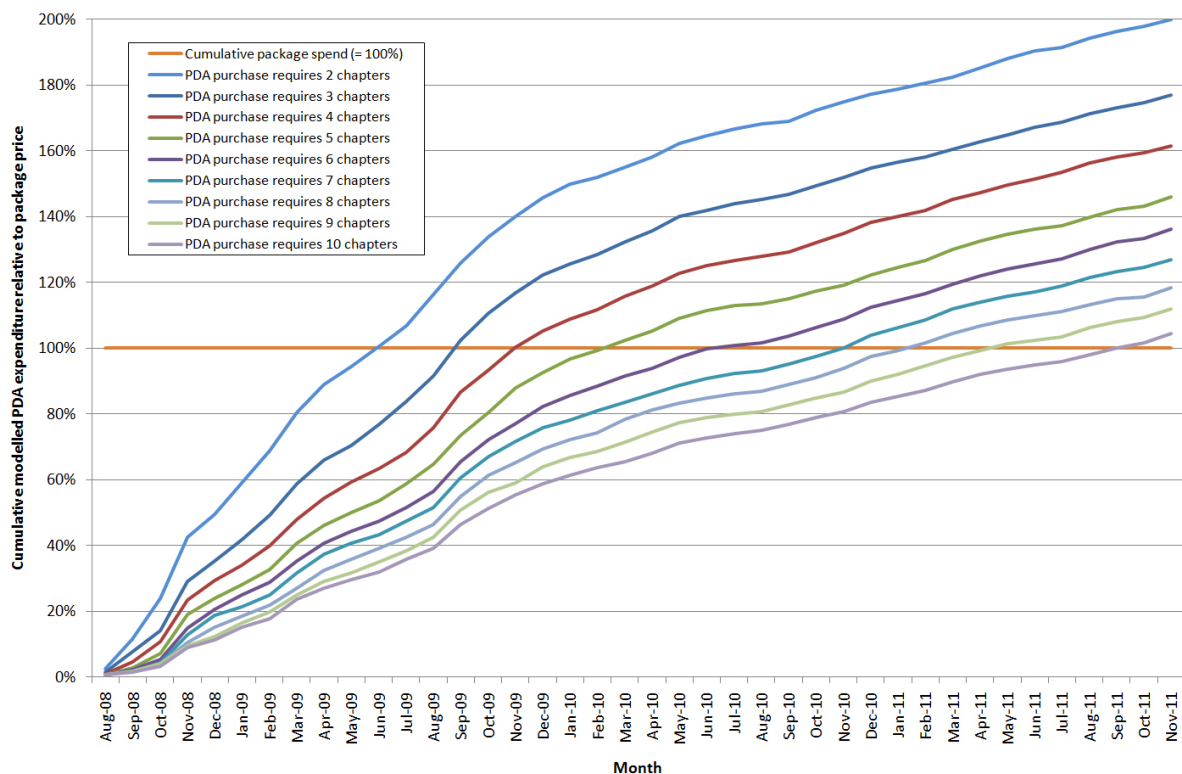


Figure 8. Modelling of Springer 2008 collection COUNTER statistics against a PDA model whereby ten page views are required to trigger a PDA purchase. The model varies the number of chapter downloads required to reach this ten-page threshold from two chapters to ten chapters.

## Evidence-based selection (EBS)

If packages do provide better value for money than PDA over the long term, the problem remains: how can the library decide which packages to buy? Ideally, the library would need to have usage statistics for a package over a year or more before deciding whether it should buy the package. But by then it is too late if the answer is no!

Evidence-based selection (EBS) offers just this sort of facility. Under this model, the library pays a relatively modest up-front fee in order to be able to access a collection for a year. Towards the end of that year, the library can evaluate its usage reports (probably with the assistance of the vendor) to decide which titles or collections to retain permanent access to, with a total value up to the fee already paid. At that point access to the non-retained titles is lost unless the library and vendor agree another year of EBS. Of course, if the library wants to spend more to retain access to more of the content, the vendor will gladly accept!

The attraction of EBS to the library is that it reduces the risk of purchasing a package which turns out to be little used. The EBS fee should be modest enough for that not to be a worry. Further, the EBS period allows the library to evaluate whether the library's best strategy would be to purchase the whole collection, subsets (by publication year or by subject, for example) or single titles, or of course a mixture of all three. For the vendor,



59 the attraction of EBS is that it gives a minimum guaranteed level of income and may lead to larger purchases that a library would not have been prepared to make without evidence of usage.

UoL experimented with three EBS trials in 2011, though two of them did not commence until March 2011 so a full year's evidence has not yet been accumulated. In the one trial that has completed, we retained all the collections in one subject area because they were consistently well used, showing a high proportion of the titles used multiple times during the year. No other collections showed a similarly clear-cut case for retention. Although some collections showed a good cost per download, on closer examination a small proportion of titles were found to be responsible for that usage (in one case all usage was from a single title) or many titles were well used but almost all of their usage was in a single month, which gave us doubts about retaining the title for the long term. As a consequence, the remainder of the titles chosen were single titles that had a low cost per download, had been used in several months of the year and were below a certain price threshold.

We thus saw the key benefits of EBS being in helping us to decide which subject collections we might purchase on an annual basis in the future, and of having low-cost, albeit temporary, access to a large collection. Only time will tell whether the single titles that we selected will continue to be well used. We hope that our selection criteria have stacked the odds in our favour, but previous experience shows that although a collection behaves fairly consistently over time, a single title's usage patterns vary greatly from one year to the next. I like to think of e-book collections as being like a liquid or gas: they have meaningful bulk properties (like temperature, pressure and volume) but at the microscopic level their behaviour is governed by the unpredictable behaviour of single molecules.

## Conclusions

The fundamental problem with building an e-book collection (in fact any library collection) is that there will always be more books published than the libraries' users might like to read than the library can afford to buy. Libraries have finite – and often diminishing – budgets, and users' needs change from one year to the next, in part because the actual users change from one year to the next, as do curricula and research interests. Sometimes a user *must* have a *specific* book, but sometimes a user needs books on a topic, and *which* book does not matter that much. The problem for the library is deciding what the most effective way is for the library to adequately satisfy most users' needs within budget.

For core textbooks and for books that are heavily borrowed in print, it makes sense to buy the e-book (if available). The library *knows* that they will be used so it makes little sense to apply a PDA model to those sorts of titles, especially if it is a model where the library could pay substantially more than list price for the e-book because of prior rentals. The problem for libraries is getting hold of reading lists to know which books to buy. PDA might get around this problem, but it might lead to too much 'noise' being bought in seeking the 'signal'.

To build a broad research collection (and plenty of student assignments are 'research' in this sense) our experience is that well-chosen collections probably provide a better value approach to satisfying most users' needs. The problem is one of how to select the collections when budgets preclude you from buying as many as you would like. Libraries without previously-purchased packages to analyse might be able to secure trial access to packages, but access needs to be for at least a year to provide sufficient evidence. Evidence-based selection might be the answer here: the fees involved should be more modest and the library might be able to bid for one-off funding for a project to gather evidence.

Some publishers/vendors will offer packages that will contain too little material that is relevant to the library to justify purchase at the price offered, but the library can negotiate on price (using usage evidence if possible). If the package remains unaffordable, then PDA might be the best solution here and it seems that publishers with more niche offerings are showing an increasing interest in offering PDA models on their own platforms.

- 60 If libraries do shift some of their book buying towards packages, then this has profound implications on how book budgets are allocated and how spending decisions are made. Just as with e-journal big deals, there are benefits to be drawn from centralizing budgets and the library taking an overview of how the total budget should be spent for the greatest benefit of all of its users.

#### References

1. Bucknell, T, Usage statistics for Big Deals: supporting library decision-making, *Learned Publishing*, 2008, 21(3), 193–199.
2. Springer, A Survey of eBook Usage and Perceptions at the University of Liverpool - University of Liverpool eBook Study: part 2: [http://www.springer.com/cda/content/document/cda\\_downloaddocument/V7671+Liverpool+White+Paper+Part2.pdf?SGWID=0-0-45-1037538-0](http://www.springer.com/cda/content/document/cda_downloaddocument/V7671+Liverpool+White+Paper+Part2.pdf?SGWID=0-0-45-1037538-0) (accessed 26 January 2012).
3. Bucknell, T, The 'big deal' approach to acquiring e-books: a usage-based study, *Serials*, 2010, 23(2), 126–134: <http://dx.doi.org/10.1629/23126> (accessed 1 February 2012).
4. Anderson, R, 31 May 2011, What Patron-Driven Acquisition (PDA) Does and Doesn't Mean: An FAQ, The Scholarly Kitchen blog: <http://scholarlykitchen.sspnet.org/2011/05/31/what-patron-driven-acquisition-pda-does-and-doesnt-mean-an-faq/> (accessed 27 January 2012).
5. Bivens-Tatum, W, 27 May 2011, PDA and the Research Library, Academic Library: On Libraries, Rhetoric, Poetry, History, & Moral Philosophy blog: [http://blogs.princeton.edu/librarian/2011/05/pda\\_and\\_the\\_research\\_library/](http://blogs.princeton.edu/librarian/2011/05/pda_and_the_research_library/) (accessed 27 January 2012).

#### Article © Terry Bucknell

Terry Bucknell, Electronic Resources Manager  
Sydney Jones Library, University of Liverpool, PO Box 123, Chatham Street, Liverpool L69 3DA, UK  
Tel: +44 (0)151 794 2692 | E-mail: [t.d.bucknell@liverpool.ac.uk](mailto:t.d.bucknell@liverpool.ac.uk)

To cite this article:

Bucknell, T, Buying by the bucketful: a comparative study of e-book acquisition strategies, *Insights*, 2012, 25(1), 51–60, doi: 10.1629/2048-7754.25.1.51

To link to this article:

<http://dx.doi.org/10.1629/2048-7754.25.1.51>