



All Faculty Publications

2010

The Short-Term Influence of Free Digital Versions of Books on Print Sales

John Hilton III

Brigham Young University - Provo, johnhiltoniii@byu.edu

David Wiley

Brigham Young University - Provo, david.wiley@gmail.com

Follow this and additional works at: <http://scholarsarchive.byu.edu/facpub>

 Part of the [Publishing Commons](#), and the [Religion Commons](#)

Original Publication Citation

John Hilton III and David Wiley. "The short-term influence of free digital versions of books on print sales." *The Journal of Electronic Publishing*, 13 (1). (2010)

BYU ScholarsArchive Citation

Hilton, John III and Wiley, David, "The Short-Term Influence of Free Digital Versions of Books on Print Sales" (2010). *All Faculty Publications*. Paper 1358.

<http://scholarsarchive.byu.edu/facpub/1358>

This Peer-Reviewed Article is brought to you for free and open access by BYU ScholarsArchive. It has been accepted for inclusion in All Faculty Publications by an authorized administrator of BYU ScholarsArchive. For more information, please contact scholarsarchive@byu.edu.

The Short-Term Influence of Free Digital Versions of Books on Print Sales

John Hilton III and David Wiley

This paper was refereed by the Journal of Electronic Publishing's peer reviewers.

Abstract

Increasingly, authors and publishers are freely distributing their books electronically to increase the visibility of their work. A vital question for those with a commercial stake in selling books is, "What happens to book sales if digital versions are given away?" We used BookScan sales data for four categories of books (a total of 41 books) for which we could identify the date when the free digital versions of the books were made available to determine whether the free version affected print sales. We analyzed the data on book sales for the eight weeks before and after the free versions were available. Three of the four categories of books had increased sales after the free books were distributed. We discuss the implications and limitations of these results.

The Short-Term Influence of Free Digital Versions of Books on Print Sales

Introduction

A growing number of authors and publishers freely distribute their books electronically to increase the visibility of their work. These books, for both academic and general audiences, cover a wide variety of genres, including technology, law, fantasy, and science fiction. Some authors claim that free digital distribution has increased the impact of their work and their reputations as authors. [1] But beyond increased exposure, a vital question for those with a commercial stake in selling books is, "What happens to book sales if digital versions are given away?"

One answer may come from the National Academies Press (NAP), which makes the text of all of its publications freely accessible. "Consequently," reported Michael Jensen, Director of Publishing Technologies at NAP, "we are very well indexed by search engines." [2] Jensen wrote that as a result of this indexing they receive many visitors, a small percentage of whom purchase books. Jensen reported that NAP's 1997 publication "Toxicologic Assessment of the Army's Zinc Cadmium Sulfide Dispersion Tests" had 11,500 online visitors in 2006. Those visitors "browsed approximately four book pages each. Of those, four bought a print book at \$45, and two bought the PDF at \$37.50. So 0.05% of the visitors to that particular book purchased it, even though they could read every page free online." [3] Thus, a nine-year-old out-of-print

publication that otherwise would likely have been inaccessible was viewed 11,000 times and purchased six times.

The Oriental Institute at the University of Chicago digitally distributes free copies of its books, and recently reported that print sales have not decreased. Specifically they noted that “[a]fter the complimentary distribution of twenty-one titles in 2008 that had for many years only been available in print, sales of these titles increased by 7% compared with the previous two years.” [\[4\]](#)

The question of how freely distributing an electronic version of a work affects print sales is difficult, if not impossible, to answer experimentally because there is no way to simultaneously release and not release free versions of a book. It is not possible to determine causation; nevertheless, the effect of free distribution on print sales is an important issue to examine.

In the present study we explored how free digital book distribution influenced book sales in the short term by examining a series of books that were released in print at one point in time, and then later released in a free digital format. Our specific question was, “Are book sales in the eight weeks following a book’s free digital release different from the eight weeks prior to this release?” Because most books have a pattern of declining sales as time goes by, our assumption was that sales would decrease slightly in the eight weeks following the free release.

Method

We followed the lead of Tim O’Reilly in using Nielsen BookScan to track the data on book sales before and after free versions were available. [\[5\]](#) BookScan tracks point-of-sales data from most major booksellers, meaning that it tracks the number of books actually sold to customers, as opposed to books sold by distributors to retailers. Notable booksellers that BookScan does not track include Wal-Mart and Sam’s Club. [\[6\]](#) In general, BookScan estimates that it tracks approximately 70% of all book sales in the United States.

Because BookScan tracks sales by week, we had to exercise some judgment in designating which weeks were “pre” and which were “post.” For example, if a free digital version was released on a Friday, some of the sales that week would be when the book was freely available and others would not be. If the release date of the free version was such that five or more days of the week fell into either a “pre” or “post” category, we assigned it to that category. In instances where the free version was released in the middle of week we did not count that week at all in our analysis; rather we tracked the eight weeks before and after the week the free version was made available. To protect BookScan’s proprietary business information, we did not link the sales figures with specific book titles in this paper.

We organized the books we studied into four different groups. The first group consisted of seven nonfiction books that had digital versions that were released at various times. The second group consisted of five science fiction/fantasy titles that had digital versions that were released at various times. The third group consisted of five science fiction/fantasy books that were released together by Random House. The fourth group consisted of 24 science fiction/fantasy books released by Tor Books. The Tor group was different from the previous three in that Tor ran a

special promotion in which they released a new book each Friday. The book was available for free download only for one week and only to those who registered for Tor's newsletter. With the other three groups, once a book was released in a free digital format it remained available, at least for several weeks, and in many cases, indefinitely.

It is important to note that some publishers, such as the National Academies Press, allow readers to view only a page at a time, and make the downloading of an entire book difficult. This was not the case with the specific books we studied. With two exceptions all of the books were available to be downloaded as entire PDF document. The two exceptions were *Cult of iPod* and *Cult of Mac*. Rather than making PDF versions of these books available to download from a static site, the author of these two books used BitTorrent to encourage the spread of the book. [\[7\]](#)

Some books were available free in digital formats beyond PDF. All of the books released by Random House were available in Stanza, an e-book format commonly used on the iPhone, Kindle, or at Scribd.com, the social publishing site that allows anyone to post a work. Several of the Tor books were made available in additional formats such as Mobipocket, a format used on some "smart" cell phones and personal digital assistants. At a minimum, all books except the two previously described were available as complete PDF downloads.

Results

Non-Fiction Books

The non-fiction books that we studied were:

- *Blown to Bits* (Hal Abelson, Ken Ledeen, and Harry Lewis)
- *Cultural Software* (Jack Balkin)
- *Cult of iPod* (Leander Kahney)
- *Cult of Mac* (Leander Kahney)
- *The Future of Ideas* (Lawrence Lessig)
- *Trigger Happy* (Steven Poole)
- *The Access Principle* (John Willinsky)

Table 1 summarizes the results of the BookScan data for these nonfiction titles.

Table 1: Sales of Nonfiction Titles

	Books sales 8 weeks before free digital release	Books sales 8 weeks following free digital release	Net difference (post sales–pre sales)
Title 1	33	46	13
Title 2	322	338	16
Title 3	6	17	11
Title 4	13	20	7
Title 5	134	86	–48
Title 6	115	139	24
Title 7	22	33	11
Total	639	662	34

Science Fiction/Fantasy Books

The science fiction/fantasy books that we studied are:

- Seaborn (Chris Howard)
- Caught Stealing (Charlie Huston)
- Magic for Beginners (Kelly Link)
- Spaceman Blues (Brian Slattery)
- The Crooked Letter (Sean Williams)

Table 2 summarizes the results of the BookScan data for these science fiction/fantasy titles.

Table 2: Sales of Fiction Titles

	Books sales 8 weeks before free digital release	Books sales 8 weeks following free digital release	Net difference (post sales–pre sales)
Title 8	45	18	–27
Title 9	16	12	–4
Title 10	252	272	20
Title 11	160	163	3
Title 12	40	181	141
Total	513	646	133

Random House Books

On March 4, 2009, Random House announced that it was releasing free digital versions of five science fiction/fantasy books. These books are:

- Assassin’s Apprentice (Robin Hobb)
- His Majesty’s Dragon (Naomi Novik)
- Blood Engines (T.A. Pratt)
- Red Mars (Kim Stanley Robinson)
- Settling Accounts: Return Engagement (Harry Turtledove)

Because March 4 was a Wednesday, and BookScan tracks sales by weeks, we did not use sales from the week of March 2–8. Thus the dates of the eight weeks prior to the release of the free versions were January 5–March 1, and the dates for the eight weeks after release were March 9–May 3. Table 3 summarizes the results of the BookScan data for these Random House titles.

Table 3: Sales of Random House Titles

	Books sales 8 weeks before free digital release	Books sales 8 weeks following free digital release	Net difference (post sales–pre sales)
Title 13	316	252	–64
Title 14	1812	2153	341
Title 15	1612	1677	65
Title 16	413	508	95
Title 17	796	817	21
Total	4949	5407	458

Tor Books

To promote an upcoming website, Tor Books began releasing a series of free e-books on February 15, 2008. To claim the free e-book, readers needed to register for a newsletter that was sent out each week announcing the location of the next free e-book. Each Friday for the successive 24 weeks a new book was released, and the previous week's book was then no longer available. The books released by Tor were:

- Mistborn (Brandon Sanderson)
- Old Man's War (John Scalzi)
- Spin (Robert Charles Wilson)
- Farthing (Jo Walton)
- The Outstretched Shadow (Mercedes Lackey & James Mallory)
- Crystal Rain (Tobias Buckell)
- Lord of the Isles (David Drake)
- Through Wolf's Eyes (Jane Lindskold)
- The Disunited States of America (Harry Turtledove)
- Reiffen's Choice (S. C. Butler)
- Sun of Suns (Karl Schroeder)
- Four and Twenty Blackbirds (Cherie Priest)
- Spirit Gate (Kate Elliot)
- Starfish (Peter Watts)
- Touch of Evil (C. T. Adams & Cathy Clamp)
- A Shadow in Summer (Daniel Abraham)

- Orphans of Chaos (John Wright)
- In the Garden of Iden (Kage Baker)
- In the Midnight Hour (Patti O’Shea)
- Battlestar Galactica (Jeffrey A. Carver)
- Flash (L. E. Modesitt, Jr.)
- Soul (Tobsha Learner)
- Darkness of the Light (Peter David)
- Three Shadows (Cyril Pedrosa)

Table 4 summarizes the results of the BookScan data for these books.

Table 4: Sales of Tor Titles

	Books sales 8 weeks before free digital release	Books sales 8 weeks following free digital release	Net difference (post sales–pre sales)
Title 18	385	260	–125
Title 19	4,178	2,561	–1,617
Title 20	2,710	2,645	–65
Title 21	1,490	1,518	28
Title 22	440	296	–144
Title 23	1,594	1,445	–149
Title 24	79	76	–3
Title 25	240	210	–30
Title 26	843	735	–108
Title 27	294	218	–76
Title 28	77	34	–43
Title 29	240	152	–88
Title 30	99	95	–4

Title 31	1,270	1,096	-174
Title 32	150	339	189
Title 33	188	136	-52
Title 34	207	250	43
Title 35	662	695	33
Title 36	137	95	-42
Title 37	136	106	-30
Title 38	100	35	-65
Title 39	75	37	-38
Title 40	6,211	2,778	-3,433
Title 41	19	744	725
Total	21,824	16,556	-5,268

Discussion

Perhaps the most significant finding of this study was the contrasting results received by Tor and the other three groups studied. With one exception, sales of the nonfiction titles increased after a free digital release, and when the sales of the books were combined, sales were up 5%. The majority of the fantasy/science fiction books that were not part of a group release also had increased sales, and as a group their sales increased 26%, largely as a result of “Title 12.” Four of the five Random House books saw sales gains after the free versions were released; in total, combined sales of those five books increased 9%. These three groups were in contrast to our initial hypothesis that book sales would decline. Although we cannot say that the free e-books caused sales to increase, a correlation exists between a free e-book and increased print sales.

The results of the Tor book sales were quite different. Only four of the twenty-four books saw increased sales during the eight weeks after the free version was made available. Two of these books (titles 32 and 41) both had releases of paperback editions that preceded the free book by only a few weeks. Thus for the majority of the “pre” weeks, a paperback version was not available. These newly released paperback versions could easily explain why the “pre” sales of these titles were less than the “post” sales.

The book with the most dramatic pre–post difference (title 40) was released just ten weeks before the free digital version was released. It is possible that what was measured with this title was the natural decline of book sales over time instead of a result of a free version being made available. But even when these three books were excluded from the analysis, combined sales of the remaining 21 books decreased 18%.

Why were the results from Tor so different from the others? This question cannot be answered with certainty. The only thing we know is that Tor’s model of making the books available for one week only and requiring registration in order to download the book was substantially different from the models used to create free versions of the other books we studied. Further research is necessary to determine if the Tor results were related to their model of free book distribution, a natural drop in sales, or if other factors account for the decreased sales.

Conclusion

The present study indicates that there is a moderate correlation between free digital books being made permanently available and short-term print sales increases. However, free digital books did not always equal increased sales. This result may be surprising, both to those who claim that when a free version is available fewer people will pay to purchase copies, as well as those who claim that free access will not harm sales. The results of the present study must be viewed with caution. Although the authors believe that free digital book distribution tends to increase print sales, this is not a universal law. The results we found cannot necessarily be generalized to other books, nor be construed to suggest causation. The timing of a free e-book’s release, the promotion it received and other factors cannot be fully accounted for. Nevertheless, we believe that this data indicates that when free e-books are offered for a relatively long period of time, without requiring registration, print sales will increase.

Although this article has focused on print sales, it should be noted that in addition to print sales, publishers and authors may have other reasons for releasing free electronic versions. As Anderson has pointed out, there are many ways to make “free” profitable. [\[8\]](#) Increasing electronic sales may be an additional motive. For example, it is possible that Kindle book sales of second and third books in a series increased dramatically when the first book was available for free. We cannot determine if this happened, because Amazon does not release Kindle book sales figures. In addition, publishers and authors may have motivations indirectly related to sales. For example, although Tor may have lost sales as a result of their free e-book promotion, the customer information harvested and the publicity gained may have been more valuable than sales they perhaps lost.

Another factor that we did not analyze was the differences in the size of the audience for the books we studied. Even within the four groups there were large differences in total sales of specific titles. Some of the fiction books had sold several hundred thousand copies, others fewer than five thousand. Future studies might examine relationships between the potential audience for a book and the impact of free digital distribution.

In addition, we did not study how free books affect the sales of other titles by an author. For example, all the books released by Random House were the first books in a series. Future

analysis needs to be done to determine whether sales of other books by an author (e.g., later books in a series) are influenced by making one of an author's works freely available.

As books increasingly become available in digital formats, the effects of free distribution may rapidly change. The explosive growth of Kindle and other e-book formats could dramatically impact how free distribution affects for-profit sales and even alter the relative importance of print sales. As the electronic publishing industry matures it will be increasingly important to research the effects of free distribution of electronic books.



John Hilton III received his M.Ed. from the Harvard Graduate School of Education and currently is a Ph.D. student in Instructional Psychology at Brigham Young University. He is interested in researching open-access issues, particularly the creation and use of open educational resources, and looking at how free digital book distribution affects print sales and the impact of books.

John can be contacted via his website at <http://johnhiltoniii.org>.



David Wiley is an associate professor of Instructional Psychology and Technology at Brigham Young University. His previous appointments include the director of the Center for Open and Sustainable Learning, a nonresident fellow of the Center for Internet and Society at Stanford Law School, a National Science Foundation-funded postdoctoral fellow, and a visiting scholar at the Open University of the Netherlands. He is also the recipient of the National Science Foundation's prestigious Young Researcher/CAREER award.

Email: david.wiley@byu.edu

Website: <http://opencontent.org>

Notes

1. James Boyle, "Text Is Free, We Make Our Money on Volume(s)," *Financial Times*, January 22, 2007, retrieved February 2, 2009, from <http://www.ft.com/cms/s/2/b46f5a58-aa2e-11db-83b0-0000779e2340.html>. Cory Doctorow, *Content: Selected Essays on Technology, Creativity, Copyright, and the Future of the Future* (San Francisco: Tachyon Publications, 2008). James Hilton and David Wiley, "Free: Why Authors Are Giving Books Away on the Internet," *Tech Trends* (in press).
2. Michael Jensen, "The Deep Niche," *The Journal of Electronic Publishing* 10, no. 2 (Spring 2007), retrieved May 13, 2009 from <http://dx.doi.org/10.3998/3336451.0010.2066>
3. Ibid.
4. The Oriental Institute, "The Electronic Publications Initiative of the Oriental Institute of the University of Chicago," April 6, 2009, retrieved May 13, 2009 from <http://oi.uchicago.edu/research/pubs/epi.html>.
5. Tim O'Reilly, "Free Downloads vs. Sales: A Publishing Case Study," *O'Reilly Radar*, June 1, 2007, retrieved February 2, 2009, from <http://radar.oreilly.com/archives/2007/06/free-downloads-vs-sales-a-publ.html>.
6. Kurt Andrews and Philip M. Napoli, "Changing Market Information Regimes: A Case Study of the Transition to the BookScan Audience Measurement System in the U.S. Book Publishing Industry," *Journal of Media Economics* 19, no. 1 (2006): 33–54.
7. "BitTorrent is a peer-to-peer file sharing protocol used for distributing large amounts of data." "BitTorrent (protocol)," *Wikipedia, The Free Encyclopedia*, [http://en.wikipedia.org/w/index.php?title=BitTorrent_\(protocol\)&oldid=311471813](http://en.wikipedia.org/w/index.php?title=BitTorrent_(protocol)&oldid=311471813) (accessed September 2, 2009).
8. Chris Anderson, *Free: The Future of a Radical Price* (New York: Hyperion, 2009).