Storming the Gatekeepers:
Digital Disintermediation in the Market for Books

JoelWaldfogel
Carlson School of Management, Department of Economics, and Law School
University of Minnesota and NBER

Imke Reimers
Department of Economics
University of Minnesota

December 21, 2012

Digitization is transforming the market for books. Lower marginal costs have reduced prices by 10-15 percent in the past four years, and digitization has given creators the ability to circumvent traditional gatekeepers and publish their work directly. The number of self-published works has grown by almost 300 percent since 2006 and now exceeds the number of traditionally published works. Given the inherent difficulty in predicting the ex post appeal of creative products at the time of investment, a growth in available new products can substantially expand the appeal of available products. While e-book data are not systematically available, we are able to document that falling prices have increased consumer surplus by $2-3 billion per year. Using bestseller lists in conjunction with title-level data on physical sales and our best estimates of e-book sales, we document that many self-published books have substantial ex post appeal to consumers. Works that began their commercial lives through self-publishing began to appear on bestseller lists in 2011 and by late 2012 such works accounted for a tenth of both bestseller listings and estimated unit sales. In romantic fiction, self-published works account for almost a third. These changes challenge the role of gatekeepers while benefiting consumers.

Introduction

Technological change has transformed content industries such as recorded music, newspapers, movies, television, and books. The recorded music was the first to face challenges from digitization with the arrival of the Napster file-sharing service in 1999. Endowed with the opportunity to obtain music files without payment, consumers withdrew from purchasing recorded music. A large body of research documents harmful effects of file-sharing on recorded music revenue, plausibly explaining all of the reduction in recorded music revenue. [1] Newspapers too have faced substantial challenges from digitization, and their revenues have fallen as sharply as the revenues to recorded music, by a third to a half since the late 1990s.

While the past decade has been challenging for participants in content industries, observers and participants have begun to appreciate that new technology brings benefits, in the forms of lower costs of production and distribution, along with threats arising from a handicapped ability to harvest revenue. An emerging body of work documents counterbalancing beneficial effects of new technology on the flow of new musical works. Waldfogel (forthcoming a,b) documents that the service flow from new music has increased since the late 1990s, along with some evidence explaining how this could be so.

The beneficial effects of new technology on the availability of new products have been obscured in both the view of researchers and in public policy debates by the negative impacts of new technology on revenue. This is understandable given the sequencing of the various effects of new technology in music. It took four years from the dawn of widespread online stealing with Napster until the availability of a viable and attractive outlet for purchasing digital music in the

[1] See, for example, Oberholzer-Gee and Strumpf (2007), Rob and Waldfogel (2006), Zentner (2006) for empirical evidence on piracy in music. Liebowitz (2011) provides evidence that file sharing explains the bulk of revenue reduction in music.

form on the iTunes Music Store. Thus, for four years, digitization appeared only to be harming music producers.

Digitization has unfolded differently in the market for books. Until the launch of the Amazon Kindle in 2007, there was no widely adopted platform for legal or illegal consumption of digital books. While one could view, say, a pdf file on a computer, the legal and illegal markets for digital books remained small prior to Kindle. Since Kindle's launch, e-readers have diffused rather rapidly. By the beginning of 2012, the share of households with a digital reader had grown to 30 percent. The US market for digital books has grown correspondingly, to 5 percent of the market for trade books in 2010, to about 15 percent in 2011, and to a quarter of the market during the first half of 2012.[2]

The evolution of the e-book market is interesting in itself as a case study of a new product. The market for e-books is also interesting as a context where digitization's possible impacts operating through cost reductions are not obscured by widespread digitally-enabled theft. Rather than piracy, digitization has had two different major impacts on the market for books. First, digital distribution has reduced the marginal cost of books to essentially zero, which has substantially reduced prices. The average retail list price for a physical fiction book was about $18 in 2008, and the average fiction e-book was priced at $9.10. Since then physical prices have been constant while average e-book prices have fallen to $5.31.[3] E-books had grown to roughly a quarter of all sales by 2012, suggesting a sales-weighted average price reduction from $18 to about $15, or about 15 percent between 2008 and 2011.

[2] See http://www.publishersweekly.com/pw/by-topic/industry-news/financial-reporting/article/53112-industry-sales-pegged-at-27-2-billion.html for data on 2010 and 2011 and http://www.digitalbookworld.com/2012/childrens-e-books-surge-in-first-half-of-2012-paperback-sales-sag/ for data on the first half of 2012.
[3] The average price of e-books does not include free books.

The low costs of e-book distribution also allow for a second effect. Making a new book meaningfully available to consumers has traditionally required the assistance of one of the six major world publishing houses, which have acted as gatekeeper of literary commerce. Online platforms such as Amazon's Kindle Direct Publishing arm, Smashwords, Lulu, and others make it possible for authors to circumvent the traditional publishing gatekeepers to make their products directly available to consumers. Self-publishing has grown substantially since 2011: the number of new self-published works now exceeds the number released traditionally. Because these products would not previously have made their way to consumers, self-published works have augmented the available choice set. This is potentially a consequential phenomenon: some of the best-selling titles of 2012 (the *Fifty Shades* series) thus far all began their commercial lives as self-published works.[4]

New technologies may therefore have brought about welfare benefit through both movements along the existing demand curve for books as well as an outward shift of the demand curve due to an increase in the appeal of product offerings. This paper seeks to describe the welfare effects of the growth in electronic books, and we proceed in seven sections.

Section 1 presents background, including the diffusion of e-readers, the growth in e-book sales overall, relative to physical books, the evolution of prices for physical and electronic books, and the evolution of the numbers of new titles available each year. Section 2 presents a simple theoretical characterization of the ways in which digitization may have changed the welfare of market participants. Section 3 describes the data available for the study. The market is new, and relatively little data is publicly available, so the data are collected from disparate and sometimes fragmentary sources.

[4] See, for example, http://www.amazon.com/best-sellers-of-2012-for-books .

Section 4 provides evidence on changes in the information environment facing consumers, showing that consumers now have more access to more information about new books, making it possible for them to find appealing products among the many newly available books.

Section 5 turns to evidence on the welfare impacts of digization. First, we document that as prices have fallen with the adoption of e-books, physical book sales have fallen and digital sales have risen, so that units sales have risen overall. This gives rise to a rudimentary estimate of the elasticity of demand for books in the aggregate and a simple estimate of the welfare benefit arising from the price reduction. We find that price reductions have given rise to a movement along the demand curve generating $5.7 billion in additional consumer surplus since 2008 and $2-3 billion per year during 2011 and 2012. Next, we turn to the supply of new products. First, we show, using both USA Today and Nielsen data on physical book sales, that the concentration of bestsellers' sales has decreased over time, suggesting that new kinds of titles are drawing market share away from existing titles. Second, we show that self-published books make up a rapidly growing share of top-selling books. As of the second quarter of 2012, books that began their commercial lives self-published made up about 10 percent of the titles. This major finding echoes findings elsewhere that works that previously would not have made their way to consumers make up a large and growing share of commercially successful works.

Section 6 uses sales ranking data described in section 3 to develop title-level estimates of the quantity of e-book sales. The resulting synthetic sales data allows us to estimate the share of sales attributable to e-books, as well as the share of sales attributable to books that were originally self-published.

Section 7 then turns to an aesthetic perspective on the new products. While it is difficult to deny that the new, self-published products add a substantial and growing component to consumer surplus, self-published work has been derided by a number of critics and traditional authors as low-quality "mommy porn."[5] More broadly, one can have the concern that large-scale entry by amateur authors would coarsen consumption patterns. This is, of course, a difficult proposition to test, but one way to assess it is to examine the overlap between commercially successful and critically acclaimed works. The New York Times produces a list of 100 "notable" books each year (available since xx). We document the degree of overlap between this list and the USA Today bestseller lists for the years that both are available. We find that approximately xx percent of bestseller sales were attributable to NYT notable books in 1995 and that this share has held steady through 2011.

We therefore conclude that self-published books offer substantial welfare gains, both through reduced prices and by making available new varieties that would not earlier have been available to consumers. We find no evidence that the availability of such products has coarsened US book consumption. The evidence in this paper adds to emerging evidence elsewhere that digitization has important benefits for both consumers and creators, even as it creates challenges for many existing intermediaries.

I. Background

1. The Traditional Publishing Industry

The publishing industry has traditionally been dominated by a handful of major publishing houses. As of mid-2012 there were six: Hachette, HarperCollins, MacMillan,

[5] See http://latimesblogs.latimes.com/jacketcopy/2012/03/bestselling-mommy-porn-50-shades-of-gray-.html or http://www.nytimes.com/2012/03/10/business/media/an-erotic-novel-50-shades-of-grey-goes-viral-with-women.html?_r=2& .

Penguin, Random House, and Simon & Shuster. Over the past decade, book retailing has become increasingly concentrated. As of mid-2012 (after the Borders bookstore chain declared bankruptcy), Amazon had 30 percent of the retail book market.[6]

The growing concentration of book retailing has led publishers to seek ways to deal more effectively with retailers, chiefly through merger. In October 2012 Penguin and Random House announced their intention to merge. The New York Times reported that the deal

> "would give the new company, to be called Penguin Random House, greater scale to deal with the challenges arising from the growth of electronic books and the power of Internet retailers. Publishers are increasingly worried about the leverage wielded by Internet giants like Google, Apple and, especially, Amazon. These companies have vast resources to invest in new technology, like digital sales platforms, and the size to let them negotiate better terms on prices."

2. Electronic Books and Costs

While e-books can be read on computers, and have therefore been in principle available for over a decade, e-books are most useful to consumers when consumed on small hand-held devices, such as e-book readers or tablet computers. The e-book market has grown quickly since 2007, driven largely by the success of e-readers from Amazon, Apple, and Barnes & Noble. While Sony had released some electronic book readers as early as 2004, the e-book market began in earnest with Amazon's release of the Kindle, priced at $399, in November 2007. The Kindle was well received by technology critics. Built in wi-fi for quick book downloading was widely applauded (Pogue, 2007). The Kindle 2 was released in February 2009, priced at $259. Later that year (November) Barnes & Noble released the Nook, priced at $259. The Apple iPad was

[6] Jim Milliot . "Amazon Picks Up Market Share: Demise of Borders, more e-books boosts the industry leader." Publishers Weekly, Jul 27, 2012. (http://www.publishersweekly.com/pw/by-topic/industry-news/financial-reporting/article/53336-amazon-picks-up-market-share.html)

introduced in April 2010, priced between $499 and $829, depending on options. In November 2010, BN released the Nook color at $249. In November of 2011, Amazon released the Kindle Fire. As of late 2011, Kindles were available at a range of prices, depending on options. For $79, one could buy an ad-supported Kindle with wi-fi. The same model was available at $109 without ads.[7]

These readers have spread quickly. As Figure 1 shows, the share of US adults owning an e-book reader grew from 2 percent in April 2009 to 19 percent in January 2012. The share owning a tablet grew from 3 percent in May 2010 to 19 percent in January 2012. The share some viable method for consuming electronic books – either a tablet or an e-reader – reached 29 percent in January 2012 (up from 18 percent before the December 2011 holiday (Rainie, 2012).

E-books have far lower costs than physical books. In 2010 the cost structure for a hardback book priced at $26 included the following components: the publisher received $13 in revenue and paid $3.90 to the author, $3.25 for printing, storage, and shipping, $0.80 for design, typesetting, and editing, and $1.00 for marketing.[8] The average per-copy cost was therefore roughly $9. By contrast, for an e-book priced at $9.99, the publisher received $6.99 in revenue. The per-copy costs the publisher faces were about $2.25 for the author, $0.38 for digitization, typesetting, and editing, and $0.60 for marketing. The average per-copy cost was roughly $3. Of course, design, typesetting, editing, and marketing are not really marginal costs, while royalties and printing are more clearly marginal costs. Even including only these plausibly marginal costs gives a marginal cost of about $7 for a physical book compared with $2 for an electronic book.

[7] See page Rainie, et. al (2012), page 15.
[8] See Motoko Rich, "Math of Publishing Meets the E-Book." New York Times. February 28, 2010.

2. Self-Publishing and Growth in New Titles

In addition to digitization's effect on book prices, digitization has also affected the supply side of the market. It is now possible for authors to make their works available direct to the public, without using a traditional publisher. Major providers of self-publishing services include Smashwords, Author Solutions, and Lulu, as well as Amazon. Amazon's Kindle Direct Publishing allows authors to sell their works through Amazon, receiving 70 percent of the sales price as a royalty (authors also pay Amazon some delivery fees). Smashwords offers a similar service, although Smashwords does not (as of November 7, 2012) distribute through Amazon.[9]

Bowker tracks the number of works released each year; they report that the number of self-published titles has grown from 61,000 titles per year in 2006 to 235,000 titles in 2011.[10] According to Bowker, "Smashwords topped the e-book producers with 40,608 titles (nearly 47 percent of total self-published e-books). The combined divisions of Author Solutions (part of Penguin Group) produced a total of 47,094 titles and Lulu Enterprises checks in with 38,005 titles." Outside of these four firms, "no company has more than 10 percent of market share," according to Bowker.

II. Theoretical Framework

Digitization, as will be detailed empirically below, has two effects on the markets for books. First, the ability to distribute books electronically reduces costs and has reduced prices as

[9] See http://www.smashwords.com/about/how_to_publish_on_smashwords.
[10] See "Self-Publishing Sees Triple-Digit Growth in Just Five Years, Says Bowker." October 24, 2012. http://www.bowker.com/en-US/aboutus/press_room/2012/pr_10242012.shtml).

well. This gives rise to a movement along the demand curve for books and possible increases in both consumer and producer surplus.

Second, digitization has also allowed creators to circumvent traditional gatekeepers – publishing houses – and make their works directly available to consumers. The number of new titles brought to market has increased substantially. If some of these are appealing to – and discoverable by – consumers, then the availability to this larger set of products can deliver welfare benefit to consumers.

Even before digitization, the book industry brought thousands of new products to market each year. Given the large number of available products, additional new products need not have added much of value to the choice set. Yet, the nature of media products makes it possible for growth in available products in have substantial benefits to consumers.

Books, like music and movies, are products whose success is difficult to predict at the time that investments are made. Caves (2000) describes the book business, along with music and movies, as industries in which "nobody knows" which products will succeed with consumers at the time that investments are made.[11] While it is far less costly to bring a new book to market than a new movie or music album, the costs of traditional publishing still limit the number of new products. Traditionally, a publisher needed to acquire rights, usually with a five-figure advance, edit and print books, promote these works to critics and, in some cases, directly to consumers. Publishers also needed to ship books to stores and, usually, to incur the costs of shipping the unsold books back from stores.[12]

[11] Screenwriter William Goldman once famously declared that in Hollywood, "nobody knows anything" about which movies will succeed. See Kenneth Turan, "What Dark horse with next 'Sunshine'?" Chicago Tribune, January 17, 2007. http://www.chicagotribune.com/topic/zap-et-sundance17jan17,0,5602793.story

[12] See Greco (2005) for detailed information about the book publishing industry.

New digital technology effectively reduces costs and allows creators to bring new work to market with enlisting the use of the publishing industry's costly apparatus. As we detail below, this has given rise to large growth in the number of new titles available. Whether new varieties benefit consumers depends largely on the predictability of books' appeal at the time of investment. If publishers and authors had perfect foresight about books' appeal to consumers, they would then release all works with (ex ante and therefore ex post) revenue in excess of costs. A decrease in the cost of releasing works would raise the number of titles released, but the additional works would, by construction, have limited appeal (with revenues below the old threshold and above the new, lower one). The benefit of title expansion would be modest. But in the more realistic case of unpredictable appeal, however, an increase in the number of titles available can substantially increase the number of titles that turn out, ex post, to have significant appeal to consumers (see Tervio (2009) for a model that embodies this logic).

In our context the question is whether the newly available works, which would previously not have been available to consumers, end up as a significant share of the commercially successful books. For empirical purposes I take the self-published works to be examples of works which would previously not have been available to consumers.[13]

The recent history of the book industry, along with this theoretical background suggests a number of question for this paper to explore. First, what has happened to the evolution of prices, units sold, and the number of book titles available over the past few years as digital technologies have diffused in the book market? Second, how has the information environment for product

[13] I realize that some self-published works are produced by authors with a history of publishing via traditional channels and who may have been able to release their self-published works through traditional channels. These appear to be a minority of cases, however. When authors can get released through traditional publishers, they typically choose that route. Many of the works in the data below originally appeared as self-published works but were picked up by publishers as they work became popular. In those cases the self-published works leaves the market, replaced by a traditionally published version.

discovery changed? Finally, have the newly available titles brought much benefit to consumers, i.e. have they had much sales success? We turn to these questions below, beginning with a discussion of available data.

III. Data

Data availability is a major obstacle to the authoritative study of the market for books in general and electronic books in particular. Ideally, we would observe the full list of available new titles from the last few years, along with title-level sales and prices for both physical and electronic books by week. The available data fall far short of this ideal but still allow some meaningful analysis. Some data area available at the aggregate level, while others are available at for individual titles over time. Some of the micro data, particular on title-level sales of e-books by week are not directly available and must be inferred from bestseller list ranks. In this section I describe the sources of data along with some of the recent patterns as prelude to more detailed analysis of the following sections.

1. Aggregate Data

We have aggregate data on US list prices and quantities sold, for physical and electronic books, for 2008-2011. We also have a few sources of estimates of the numbers of new titles of physical and electronic books over the same period.

Data on top-level quantities sold are from BookStats, which calculates total sales based on surveys of a large number of publishers.[14] As Table 1 shows, they report 2008 sales of 2.16 billion trade books, a category that includes adult fiction, adult nonfiction, juvenile books, and religious books. (The major categories not included in trade books are professional and educational books). Of these sales, the majority (68 percent) were various forms of softcover

[14] See http://bookstats.org/bookstats-2012.php as well as many press accounts summarizing the statistics.

books (trade and mass market paperbacks), while just over a quarter (26 percent) were hardcover books. E-books made up a very small share (0.4 percent). Since 2008, trade paperbacks and hardcover books have held relatively steady. By contrast, sales of mass market paperbacks (the inexpensive paperback format printed on low-quality paper and designed to be sold on racks in airports, etc.) have fallen, and sales of e-books have risen sharply. In 2011 total trade units sold stood at 2.50 billion, an increase of 10.7 percent in units sold over 2010. Electronic books accounted for 15.5 percent of the trade units sold in 2011.

Full-year data for 2012 are not yet available, but as of August the Association of American Publishers has released year-to-date data on the change in category revenue relative to 2011.[15] Relative to 2011 year-to-date revenue in adult hardbound books was down by 1.6 percent, up by 10.5 percent in softcover trade, down in mass market paperbacks by 16.3 percent, and up in e-books by 36.9 percent. Applying these revenue percentage changes to the unit totals for 2011 gives rough estimates of 2012 quantities. E-book units are thus estimated at 531 million, while combined physical units are estimated to be 2.17 billion, putting total units at 2.70 billion.[16] This is an increase of 8 percent over 2011 unit sales.

We have aggregate data on retail prices from a few different sources. One systematic source is the Library and Book Trade Almanac 2012 (Bogart 2012), which reports data on the number of new titles by format and subject along with their retail prices. The Almanac obtains these data from book wholesaler Baker & Taylor. The prices they report are the simple averages across titles; they are not weighted by sales of books. Table 2 reports these prices (in 2011 dollars, adjusted using the CPI) for overall titles as well as for fiction alone, by books formats. A

[15] See the August :StatShot" in http://www.publishersweekly.com/pw/by-topic/industry-news/financial-reporting/article/55082-aap-monthly-statshot-august.html .

[16] We weight the percentage change for distinct physical categories (hardbound, softcover trade, mass market) using their respective 2010 shares of physical sales. These imply 2.7 percent increase in physical units 2011 to 2012.

few patterns are clear. First, especially for fiction, mass market paperbacks and e-books are substantially less expensive than other formats. Second, physical book prices appear to be relatively constant, while e-book prices are falling. For example, the average fiction e-book had a retail price of $9.10 in 2008 and $5.31 in 2011, a reduction of 42 percent.

There is fragmentary data on e-book prices available from other sources. A firm called iobyte solutions collects data on the prices of bestselling books, and they report the the distribution of prices over the past 6 months by cells ($0-2.99, $3-7.99, $8-9.99, $10 and over). Figure 2 reports the share of Kindle top 100 Bestsellers priced below $8 and $10, respectively, between April and December 2012. The share priced below $10 has risen from 60 percent in late April to 80 percent in late December. The share priced below $8 has risen even more sharply: from 40 percent in late April to nearly 65 percent in late December. The iobyte evidence corroborates that declining list prices for electronic books resultd in lower prices paid for books with substantial sales during 2012.

The Bureau of Labor Statistics CPI provides another source of information on book prices. The BLS collects data on the prices of recreational books. These data appear to include electronic books. The instructions for interviewers include two book categories, "books purchased through book clubs" and "other books, audio books, or e-books."[17] The latter category includes "all new books, audio books, and downloadable e-books purchased individually." The price series for recreational books for all urban consumers is displayed in Figure 3, 2002-2012. This series is not seasonally adjusted, and December 1997=100. The price level for books is steady at 104 between 2002 and 2007. It then rises from 2007 until the end of 2009, reaching a peak at 107. Since the beginning of 2010, the series has been falling fairly steadily, reaching 100 in late 2012. In the past two years, the price of books has fallen by 6.9 percent according to the

[17] See http://www.bls.gov/respondents/cpi/tpops/group_13.htm#books .

CPI. Given that the overall CPI rose by nearly 9 percent between late 2009 and late 2012, the price of books has fallen 14.6 percent in real terms over this period.

It seems clear that book prices are falling, but list prices (such as those in Table 2) likely overstate actual prices paid for books. For an estimate of current price levels for physical and electronic books we monitored Amazon's top 100 bestselling physical and electronic books during November and December 2012. Weighting prices by a rough estimate of sales (the reciprocal of the sales rank), the average e-book price was $7.95, and the average physical book price was $13.83. During 2012 e-books have made up 19.6 percent of units sold (using our estimates in Table 1), so the average price paid for a book during 2012 is 0.196(7.95) + (1-0.196)(13.83) = $12.68. Using the annual average of the real CPI for recreational books (books CPI/overall CPI), we can construct a series of book price levels. The average 2012-dollar price of a book by this approach was $14.13 in 2008, $14.31 in 2009, $13.93 in 2010, $13.19 in 2011, and $12.68 in 2012.

Table 3a reports the number of new titles available in each year, overall and for fiction in particular. These data are drawn from Bogart (2012). The number of new physical titles overall was 180,032 in 2008, with 18,638 in fiction. Since 2008, the number of new physical titles has been essentially constant: total titles numbered 177,126 in 2011, while physical fiction titles numbered 19,760. The number of electronic titles, by contrast, has increased substantially. Total new electronic titles grew from 35,495 in 2008 to 111,150 in 2011. For fiction alone, the number of new electronic titles grew from 7,414 in 2008 to 39,886 in 2011.

Bowker, the publisher of the Books in Print database, also produces an annual time series on new titles, which we report in Table 3b. Their overall time series does not distinguish physical from electronic books. According to Bowker, the number of new titles grew from

215,138 in 2002 to 347,278 in 2011. Fiction alone increased from 25,102 to 60,075 over the same period. Bowker also produces counts of the number of self-published works 2006-2011, distinguishing physical from electronic.[18] In 2006 there were 60,875 new self-published works released, roughly 90 percent of them physical. The number of new self-published physical works has tripled, reaching 148,424 in 2011. The number of self-published electronic works has grown by more than 10 times, from 7,758 in 2006 to 87,201 in 2011.

Alone and together, these data sources indicate a substantial growth in the number of new titles available each year since 2006. The increase in the number of new works has been driven entirely by self-published books; and many of these are electronic.

2. Title-level information on physical books

The aggregate data are useful for making a rudimentary calculation of the welfare benefit from lower prices, but without information on sales volume of new self-published titles we cannot say much about the benefit that they produce for consumers. To remedy this, we attempt to construct a title-level dataset including sales of each title in physical and electronic forms, as well as an indicator of whether the work was originally self-published.

Nielsen BookScan collects title-level data on sales of trade books, and they have these data back to 2003. BookScan has two shortcomings. First, Nielsen allows only-title-by-title access to their data. Given the time it takes to obtain a single title, it is difficult to obtain the universe of literally hundreds of thousands of titles published per year. Second, Nielsen does not include e-books in their data.

[18] See Bowker, Self-Publishing in the United States, 2006-2011: Print vs. Ebook http://www.bookconsumer.com/store/product.php?id=37

While these shortcomings limit what one can do with Nielsen data, they have some offsetting benefits. Mainly, Nielsen does compile weekly bestseller lists, which include the 100 bestselling trade titles that week, along with the weekly sales of each title. These are available back to 2003. Each entry includes both the quantity sold and the retail list price. The Nielsen data also include the publisher and imprint, along with the ISBN number. The data do include self-published books but only if they have physical sales. Aggregating the Nielsen weekly top 100 gives 94.9 million units sold in 2009, a year in which the total sales of physical trade books was 2.2 billion. By 2011, when total physical sales were 2.1 billion, the Nielsen weekly top 100 books accounted for 72.2 million.

3. Title-level information on electronic books

We are aware of no systematic direct source of weekly information on the title-level sales of electronic books. That said, Publisher's Weekly (PW) has released lists of the top-selling electronic book titles for 2010 and 2011. For 2010 they included titles selling more than 10,000 copies; for 2011, they included titles selling over 25,000. Their list does in principle include self-published books, but it excludes titles priced below $5.00. De facto, this excludes many self-published titles. PW includes 30 adult, and 82 children's title electronic sales for 2010. PW's 2011 list included 329 adult titles. Electronic sales of these titles total 6.6 million and 26.6 million for 2010 and 2011, respectively. In about a quarter of cases, PW reports only a sales rank rather than a quantity. We can impute a quantity by regressing log q on the book's log rank, where the rank is within year and type (adult vs children). Using this imputation we observe title-level electronic book sales of 7.2 million in 2010 and 33.8 million in 2011.

Bestseller lists do convey some information about the sales of both electronic and physical books. USA Today produces a single weekly list based on overall physical and electronic sales, and self-published books are included. In addition to producing a ranking of the top 150-selling books each week, the USA Today list indicates the format (hardback, paperback, or electronic) selling the most copies of the title this week. The USA Today list, available back to 1993, includes over 1,000 separate titles per year. The list includes the name of the publisher, which is helpful for determining whether a book is self-published.

Figure 4 shows the share of listings for which the electronic version was the bestselling edition, by week. The share was essentially zero prior to 2010 and rose to about 3 percent during 2010. The share jumped markedly at the start of 2011, most likely because of the heavy volume of e-readers given as gifts at Christmas during 2010. The share was roughly steady at 25-30 percent during 2011, then fell toward the end of the year, presumably because e-books are not popular as holiday gifts. The e-book share then jumped to 70 percent at the start of 2012, again because recipients of e-book readers made their first e-book purchases immediately after Christmas. During 2012, the share of listings selling best as e-books has fluctuated between 30 and 40 percent. This figure suggests that e-book sales are continuing to grow, although the growth between 2010 and 2011 appears larger than the growth between 2011 and 2012.

Determining which books are self-published requires some detective work. Some are easy. We deem a work self-published if the listed publisher contains the word "self" (e.g. "self-published via Amazon"). We also include works published by the major self-publishing services listed in Bowker (2012). These include Smashwords, Lulu Enterprises, and various divisions of Author Solutions (Xlibris, Authorhouse, IUniverse, and Trafford). These services collectively account for about three quarters of self-published electronic books. We also found works that

had originally been self published using online sources.[19] Table A1 lists all of the works that we identify as originally self-published.

Title-level price data for electronic books are not readily available. It is of course easy to get the prices of current bestsellers at the Amazon Kindle Store; and there are some collections of average e-book prices for the last few years. But there is no obvious way to get the prices of particular e-book titles from the past. While Amazon provides an annual bestseller list archive that preserves the sales ranking order, the prices posted on the list are current prices. If a book had been self-published in the year of the list but is not available through a major publisher, the list includes no price information.

IV. The New Information Environment

Consumers have traditionally learned about new books through book reviews authored by professional critics and published in established media outlets. The number of such book reviews has traditionally been small in comparison with the number of books released each year. In 2010, when roughly 200,000 books were released, traditional media outlets published 50,000 book reviews. The largest source of reviews are publications aimed at bookstores and libraries (Publishers Weekly, Library Journal, and Kirkus). See Table 4. Because many works are reviewed by multiple outlets, the number of reviews exceeds the number of works reviewed. Media outlets aimed directly at consumers issued far fewer book reviews. For example, the New York Times reviewed about 1,250 books per year, and the Washington Post reviewed roughly 1,000. While it is difficult to say how many works were reviewed, it is clear that only a small share of works released were reviewed.

[19] See http://jakonrath.blogspot.com/2011/01/guest-post-by-robin-sullivan.html .

The past few years have seen the development of other sources of information about books. These include crowd-sourced information, notably reviews and star ratings posted at retailers themselves (e.g. Amazon) as well as reviews and ratings posted at third party sites such as Goodreads, AllReaders, or BookPage.

Of the crowd-sourced sites, Goodreads is the largest. According to Alexa.com, Goodreads is the 167th-ranked site in the US, with over 10 million users.[20] Their book coverage is broad: "Goodreads has 10 million reviews across 700,000 titles - one of the largest and deepest collection of quality book reviews on the internet."[21] Goodreads was launched in 2007 and has therefore generated reviews for over 100,000 works per year since its launch. This represents substantially more book reviews than the traditional media sector produced over time.[22]

The new information environment also includes bloggers and small-scale organizations that review books online. One can find lists of these independent reviewers in online guides to self-publishing.

According to the Step-by-Step Guide to Self-Publishing, [23]

> "Book reviews are the best way to promote your book for free! A regular print ad in a newspaper can cost thousands of dollars. A book review only costs you the postage and the cost of the book. Got it?
>
> Although submitting books for review can be costly (with postage costs, etc), the rewards are enormous. And it's still a LOT cheaper than print ads. A positive book review on Amazon is like a shower of gold for writers.

The development of new information sources, along with the growth in new products, raises the possibility that consumers will be able to find a better set of books than would have

[20] See September 4, 2012, *The Economist*, The World's Biggest Book Club.
[21] See http://www.goodreads.com/api.
[22] See See http://www.goodreads.com/blog/show/372-anatomy-of-book-discovery-a-case-study for an interesting case study about Goodreads.
[23] See http://www.stepbystepselfpublishing.net/reviewer-list.html. See also http://www.stepbystepselfpublishing.net/trick-your-way-into-getting-book-reviews.html .

been possible when the informed choice set was smaller. Whether this is so is an empirical question.

V. Digitization and Welfare

1. Surplus for Market Participants

Cost reductions and reduced entry barriers can give rise to three sorts of changes. First. Price reductions can raise consumer surplus. Cost reductions in conjunction with price changes may change producer surplus. Finally, entry of new products will consume some additional fixed costs. Here we attempt to tally these welfare changes.

How much have consumers benefitted from e-books? To a first approximation we can view the past few years of price reductions driven by new cost-reducing technology as a movement along the demand curve. In 2008, the average price of a book was $14.13, and consumers purchased $2.164 billion books. In 2012, the average price was $12.68, and consumers purchased 2.707 billion books. With a linear approximation to demand, consumer surplus in 2012 is $3.5 billion higher than in 2008. This is 10.2 percent of 2012 revenue. Since 2008, the cumulative increase in consumer surplus, relative to the 2008 level, is $5.7 billion. See Table 5.

What are the effects on producers? The marginal cost of a physical book is essentially $7, while the marginal cost of an e-book is essentially $2 (this calculation treats royalties paid to authors as a component of marginal cost). To a first approximation, then, the gain in producer surplus from is the $(Q_{\text{e-books sold}})*(5 - (P_{\text{physical}} - P_{\text{e-book}}))$. Based on the Amazon prices cited

above, this is Q*(5 – (13.83-7.95)) = Q*(5 – 5.88) = -0.88*Q. In 2012, this is a reduction of $0.47 billion.

Finally, what has happened to the fixed costs of creation? Between 2007 and 2011 the number of new titles released annually has grown by about 75,000. The additional cost is the cost of producing 75,000 new books, which consists largely of the time spent researching and writing the books. For a conservative estimate of the time required to produce a book we can look at a prolific writer. Stephen King has written many novels and a large number of short stories over his career. His first book, *Carrie*, was published in 1974, and he has written steadily since then. Figure 5 shows King's cumulative page output over time. King produces about 800 pages of published fiction, or about 2 novels, per year. Given King's legendary prolific nature, using his productivity gives a conservative estimate of time cost, of 6 months of full time work per novel. It's not clear what wage that aspiring writers might otherwise have earned. If we use the average hourly pay of $23.50, it costs society an average of $27,000 per work.[24] If we instead use the federal minimum wage of $7.25, a work costs $7,250 to create.[25] The additional 75,000 works thus cost society between $544 million and $2.025 billion per year.[26]

The gains to consumers appear to exceed the costs to other parties.

2. Costs of Coming to Market and Product Selection

We have seen above that there has been a substantial growth in the number of new products available to consumers, driven largely by self-published books. Has the availability of these new products attracted consumption? We address this question in two parts. A first

[24] See "Economy at a Glance" (http://www.bls.gov/eag/eag.us.htm, accessed December 18, 2012).
[25] See http://www.dol.gov/whd/state/stateMinWageHis.htm#.UNGtAram4lg for information on the US federal minimum wage.
[26] Of course, it's not clear that the newly available works are also newly produced. It's possible that this number of works was always written without seeing the light of day.

question is simply whether the sales concentration of books has declined. If the new products are sufficiently appealing to attract consumption away from traditional products, then we should see a decline in sales concentration.

We can construct a rudimentary measure of sales concentration directly from the USA Today ranking. Each week the ranking includes 150 books; the list therefore includes 7,800 (52 x 150) entries over the course of the year. Figure 6 shows the number of distinct works on the list each year, which provides a measure of sales concentration. Between 1994 and 2000 the number hovered between 1,000 and 1,100. Between 2001 and 2005 the number fluctuated between 1,100 and 1,200. Since 2005 the number has increased more quickly and reached 1,400 in 2011 (the last full year of data). The decrease in sales concentration is consistent with the availability of new products drawing consumption away from existing products.

A second and more direct measure of whether the new products are appealing to consumers is simply the share of bestseller list entries accounted for by works that were originally self-published, displayed in Figure 7. Zero prior to 2011, the share rose to 4 percent by mid-2011, fell to 2 percent, then rose above 6 percent by the end of the year. In 2012 the share has continued to fluctuate but has reached a higher peak of 10 percent mid-year. Self-published works have had their largest impact in the romance category. Figure 8 shows the self-published share of listings among romance books. The share reaches 20 percent during 2011 and 30 percent during 2012. It is clear that self-published works, which previously would not have meaningfully made their way to consumers, have rapidly become a significant share of total sales. That self-published books make up a large share of titles is suggestive that self-published works also make up a large share of consumption, but drawing that inference requires data on sales quantities of both physical and electronic books by title.

VI. Estimating Title-Level E-book Sales

a. Estimation using USA Today Ranks

We observe title-level weekly physical sales for bestsellers. The only title-level quantity data we observe for e-books is the annual totals for bestsellers reported by Publishers Weekly. In this section we use these data to develop title-level estimates of e-book sales for all works on the USA-Today lists during 2010-2012. We do this in two parts. First we need to rank overall sales of all of the books on the USA Today list. Second, given this ranking, we demonstrate and exploit the tight relationships between overall annual sales ranks and both e-book and physical book sales. We then impute sales for those titles where we do not observe edition sales directly.

Over 1,000 titles appear on the USA Today weekly bestseller list during each year. Turning this list into an annual ranking requires a method for translating weekly sales ranks into sales quantities that can be aggregated across weeks to the year. Fortunately, have an annual top 100 sales ranking from USA Today for each year 2007-2011, and we can use the annual ranking to infer the sales-rank relationship in the weekly data. Define r_{it} as title i's ranking in a week t's data. Define r_{iy} as the title's ranking in yearly data. If $q_{it}=\alpha r_{it}^{\beta}+\epsilon_{it}$, then $q_{iy} = \sum_{t=1}^{52} q_{it} = \sum_{t=1}^{52}(\alpha r_{it}^{\beta} + \epsilon_{it})$. If we use the index i to denote annual rank, then it should of course be true that books with lower rank numbers outsell books with higher rank numbers. That is, the following inequalities should hold: $q_{1y} > q_{2y} > \ldots > q_{100y}$. The information contained in these inequalities can allow us to infer the weekly sales-rank relationship. That is, we can choose β to minimize deviations from the inequality condition. The minimand is

$$\sum_{k=2}^{100} pos([\sum_{t=1}^{52}(\alpha {r_{kt}}^{\beta} + \epsilon_{kt}) - \sum_{t=1}^{52}(\alpha {r_{k-1,t}}^{\beta} + \epsilon_{k-1,t})]$$

The coefficient α is not identified because the rank of q and q/α is the same. Using this approach, we find that β=0.76652 for 2010 and β=0.90156 for 2011. Because we observe the annual top 100 ranking, we can validate our estimate. For 2010, the correlation between the implied and actual annual top 100 is 0.9611. For 2012 it is 0.8743. The estimate of β allows us to create a full annual sales ranking from the USA Today weekly sales ranks. This gives us estimates of the sales ranks for the top 1000 books of each year.

From the Nielsen data we also observe the quantity of annual physical sales for nearly 1,000 works each year, as well as the quantity of electronic sales for a smaller subset from PW. How well do overall annual ranks track annual sales for the titles? Figures 9 and 10 show the relationships between log sales and log annual ranks for physical and electronic sales, for 2010 and 2011. We are missing e-book sales data for the majority of titles, but the strong relationship between rank and sales suggests that we can estimate sales with ranks. For 2010, we regress log e-book sales on the combined rank, and replace the missing e-book sales data with the exponentiated prediction.

Because we observe e-book sales only when they exceed 25,000, we must deal with the censoring. To this end I ascertain the sales rank relationship for e-books using an interval regression. For e-books that were not on the PW list, I assume that their sales fell short of 25,000 in 2011 and short of 10,000 in 2010. E-book sales are continuning to grow in 2012, but we do not have total end-of-year sales for the bestselling titles of 2012. Because e-book sales are continuing to rise, we can conservatively assume that the 2011 rank-e-book sales relationship

provides a conservative estimate of true e-book sales for 2012. Table 6 reports estimation results.

This process of imputation gives us a title-level dataset with observed sales of physical books, estimated sales of electronic books, and an indicator for whether the work was originally self-published, for all of the works appearing in the USA Today weekly bestseller. Note that the e-book imputation technique is only employed for 2010-2012. (This may not be a problem since e-books made up only about a percent of unit sales in 2009).

Table 7 reports estimates of sales in the USA Today sample. We estimate that total (physical plus electronic) sales of these titles was 98.7 million in 2009, 95.2 million in 2010, and 119.0 million in 2011. Of these totals, e-book sales made up 6.2 million units in 2010 and 40.3 million in 2011. For 2012 up to the end of November we estimate 78.5 million in total sales, of which 16.9 million units are electronic books.

Self-published works make up a negligible share of unit sales in 2009 and 2010, less than 0.1 percent of unit sales. Self-published works grew substantially in 2011, to 2.1 million units, which is 2.1 percent of physical sales and 1.2 percent of electronic sales. Year-to-date in 2012 self-published works account for a much larger share of sales. Of 78.5 million units sold thus far, 15.2 million units are for works originally self-published. Self-published works make up 20.9 percent of physical sales and 13.7 percent of electronic book unit sales.

VI. The Coarsening of Product Offerings and Consumption

The Fifty Shades trilogy includes by far the best-selling self-published books. While they have sold extremely well, the work is not respected by critics. For example Jen Doll, writing in the Atlantic Wire wrote,

> "Look, I'm not afraid to say it: *50 Shades of Grey* is a terrible book. I know this because I have started reading it. It didn't take long to figure out. The writing is stilted and relies on tropes that anyone who's ever sat through 15 minutes of a high school writing workshop would know to avoid. The characters are two-dimensional and stereotypical."[27]

Others have criticized the quality of self-published works. Bestselling author Sue Grafton raised a firestorm of controversy during 2012, saying in an interview that,

> "To me, it seems disrespectful…that a 'wannabe' assumes it's all so easy s/he can put out a 'published novel' without bothering to read, study, or do the research. … Self-publishing is a short cut and I don't believe in short cuts when it comes to the arts. I compare self-publishing to a student managing to conquer Five Easy Pieces on the piano and then wondering if s/he's ready to be booked into Carnegie Hall."[28]

These concerns raise the question of whether the availability of work that critics view as low-quality has coarsened consumption patterns. One way to address this is to quantify the overlap between the bestseller list and the books deemed by critics to be of high quality. The New York Times produces a list of 100 "notable books" each year. We can ask how many of these overlap the USA Today bestsellers as well as how the overlap varies over time.

Figure 11 shows the number of NYT Notable works from each year that also appear among the USA Today bestsellers at some point during the year. About 40 of the 100 NYT notable works typically appear among the USA Today bestsellers. Figure 12 approximates the USA Today works' sales as 1/(weekly rank). Using this approach, we can see that NYT notable works account for about 5 percent of the estimated sales of USA Today bestsellers. There

[27] See Jen Doll, "The Alleged Sexiness of '50 Shades of Grey'" The Atlantic Wire, May 22, 2012. http://www.theatlanticwire.com/entertainment/2012/05/alleged-sexiness-50-shades-grey/52667/

[28] See http://www.forbes.com/sites/davidvinjamuri/2012/08/15/publishing-is-broken-were-drowning-in-indie-books-and-thats-a-good-thing/ .

appears to be a downward trend from a high of 13 percent in 2000 to 3 percent in 2011, but it is difficult to determine whether the trend is statistically meaningful.

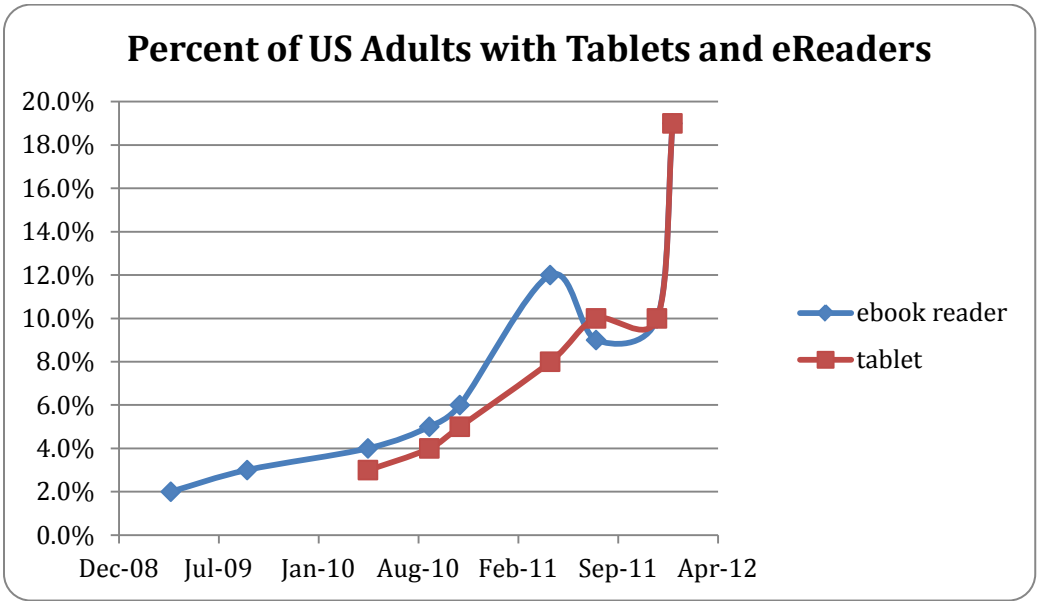
Conclusion

New technology has reduced the costs of creating and distributing books and has given creators the opportunity to circumvent the traditional gatekeepers of publishing. Price reductions have spurred book consumption, with benefits to consumers on the order of $3.5 billion per year . Self-published works, along with new institutions for product discovery, have expanded the choice set available to consumers. As of mid-2012 self published works account for about 10 percent of bestselling titles and sales.

Disintermediation presents a challenge to traditional publishers and retailers. Amazon has responded by becoming a major facilitator and retailer for self-published titles. Traditional publishers are also responding. Penguin (a division of Pearson publishing) purchased one of the largest self-publishing companies Author Solutions in July of 2012 for $116 million.[29] Coming years promise to bring challenges to the traditional industry along with benefits to consumers.

[29]See Paul Sonne and Jeffrey A. Trachtenberg. Penguin Group Dives into Self-Publishing. Wall Street Journal. July 19, 2012. http://online.wsj.com/article/SB10000872396390444464304577537092288601370.html

Figure 1



Source: http://libraries.pewinternet.org/2012/01/23/tablet-and-e-book-reader-ownership-nearly-double-over-the-holiday-gift-giving-period/

Figure 2

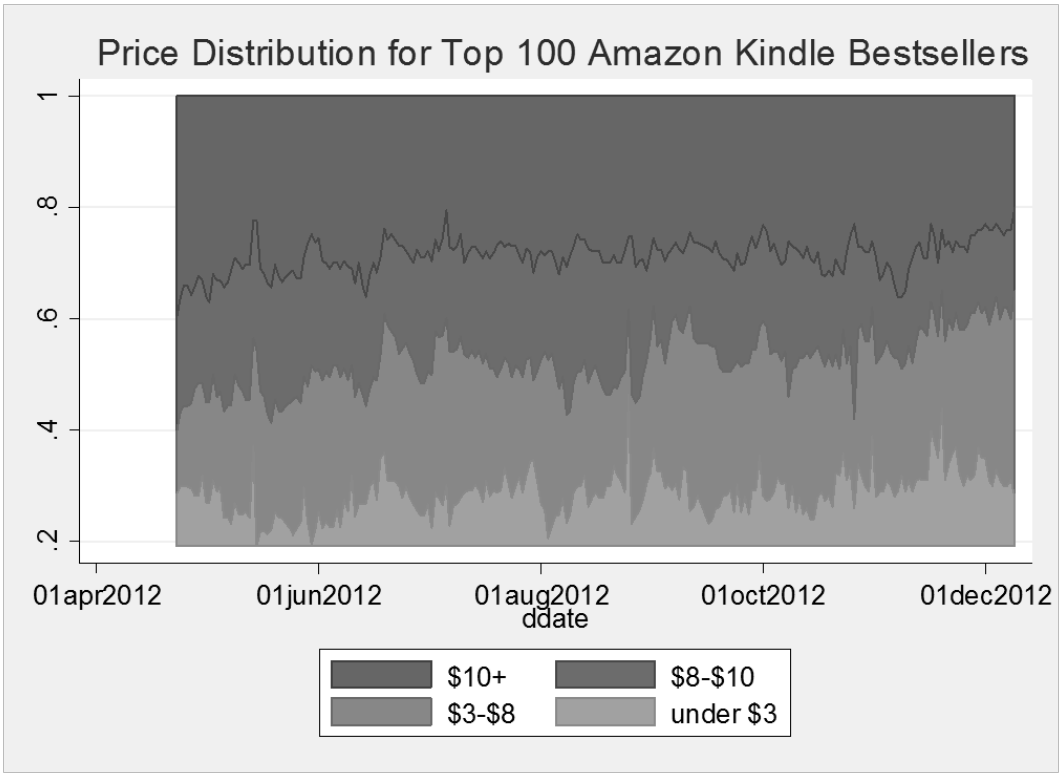


Figure 3

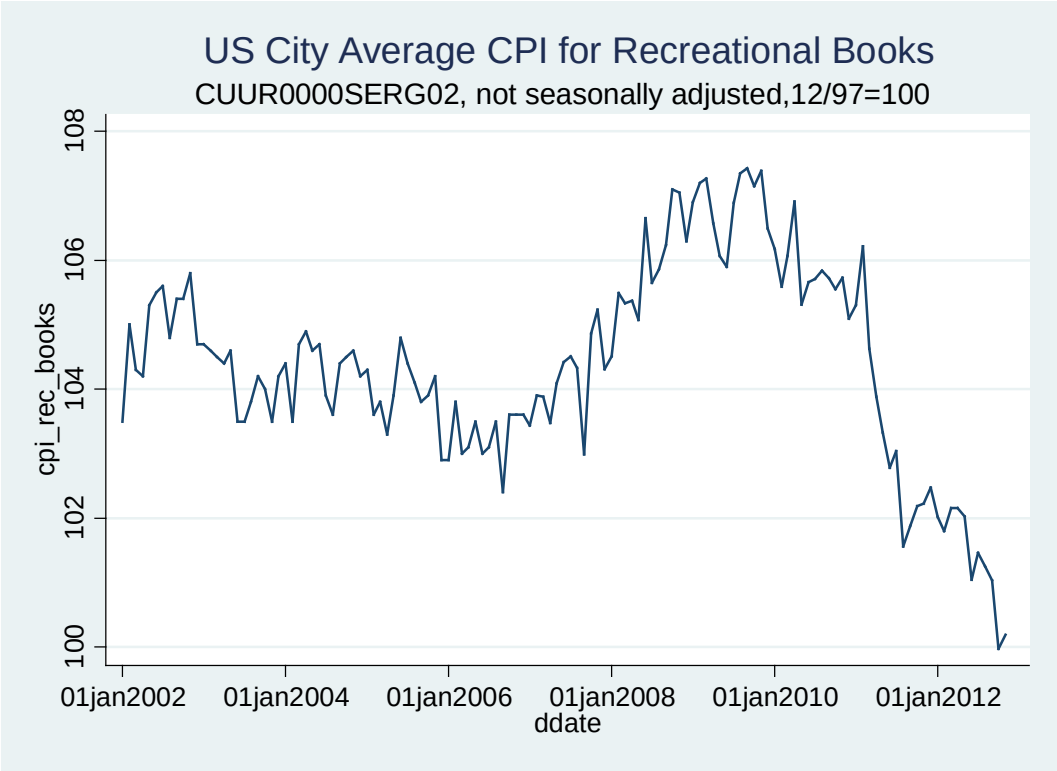


Figure 4

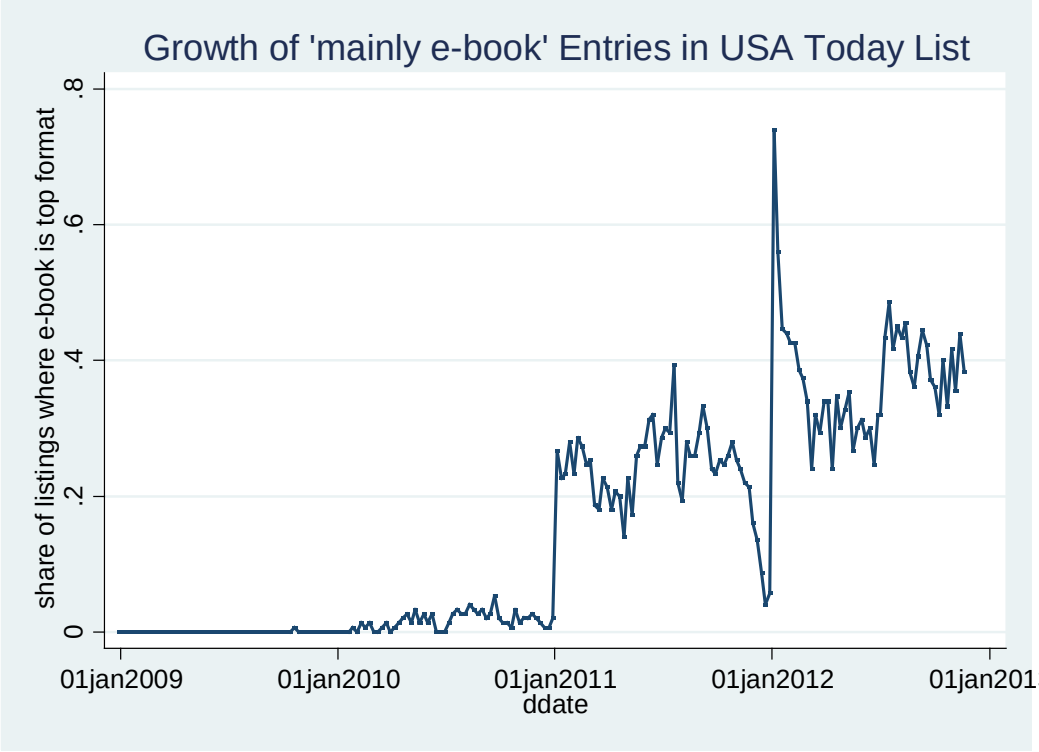


Figure 5

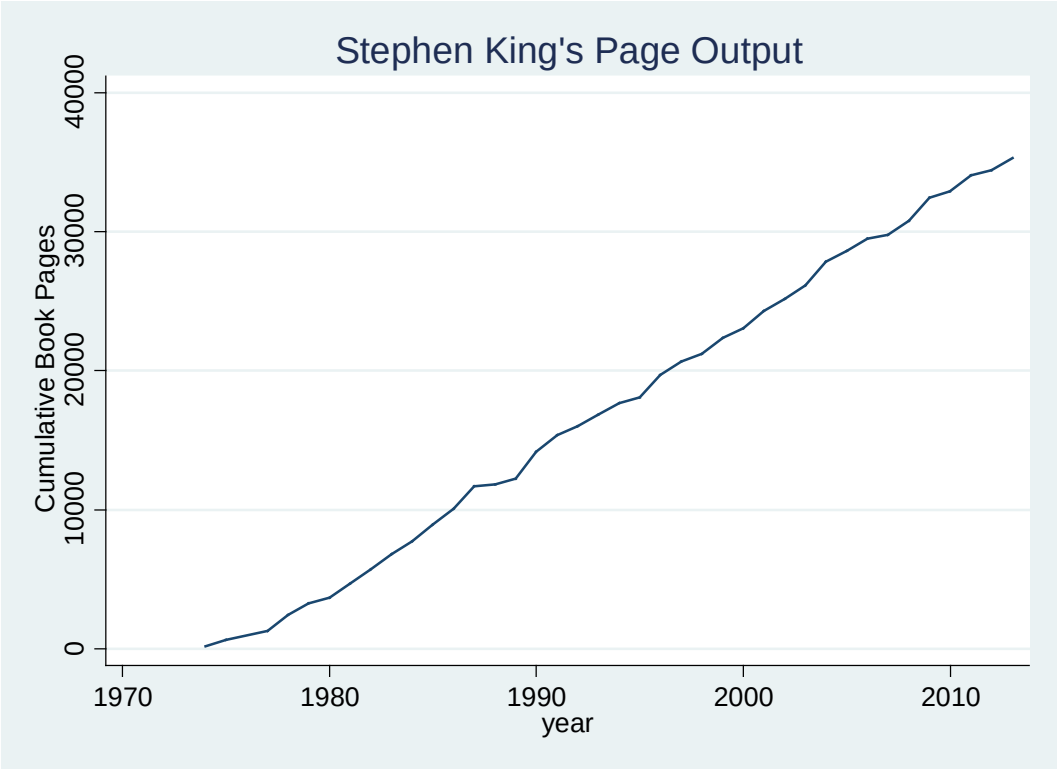


Figure 6

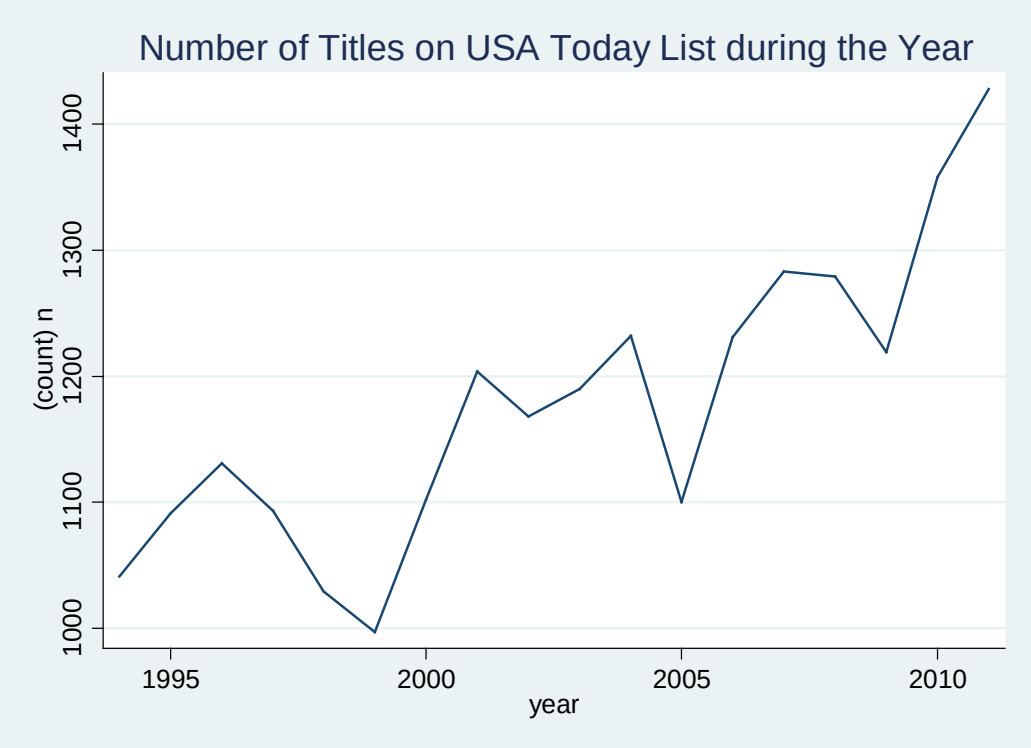


Figure 7

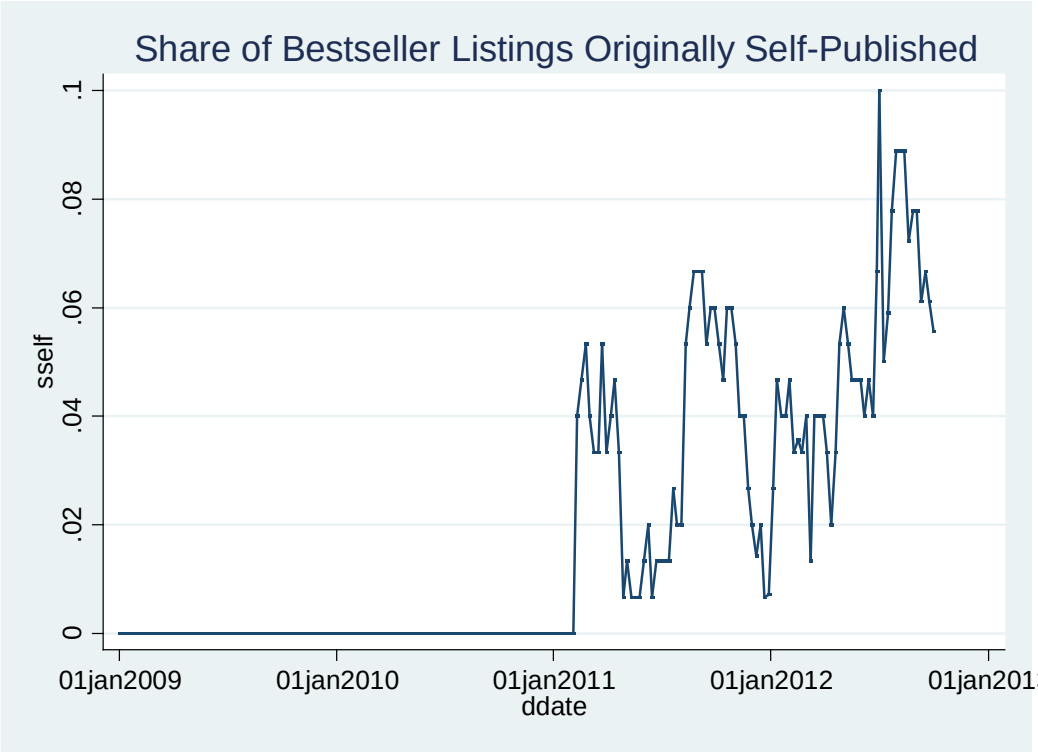


Figure 8

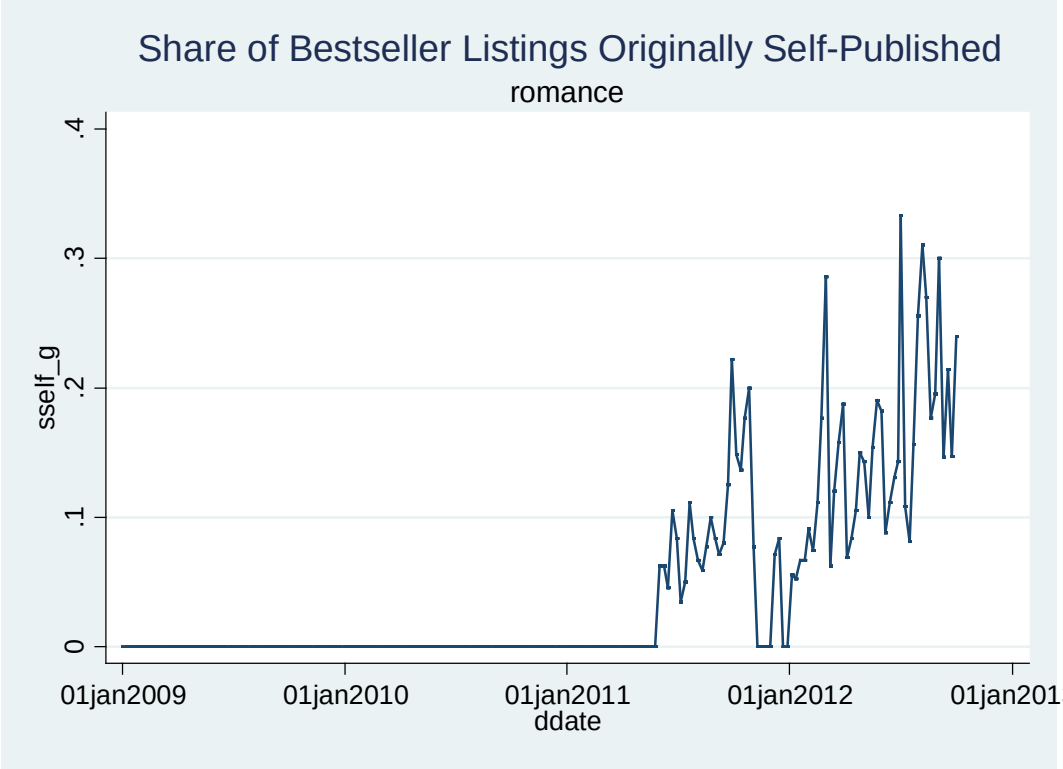


Figure 9

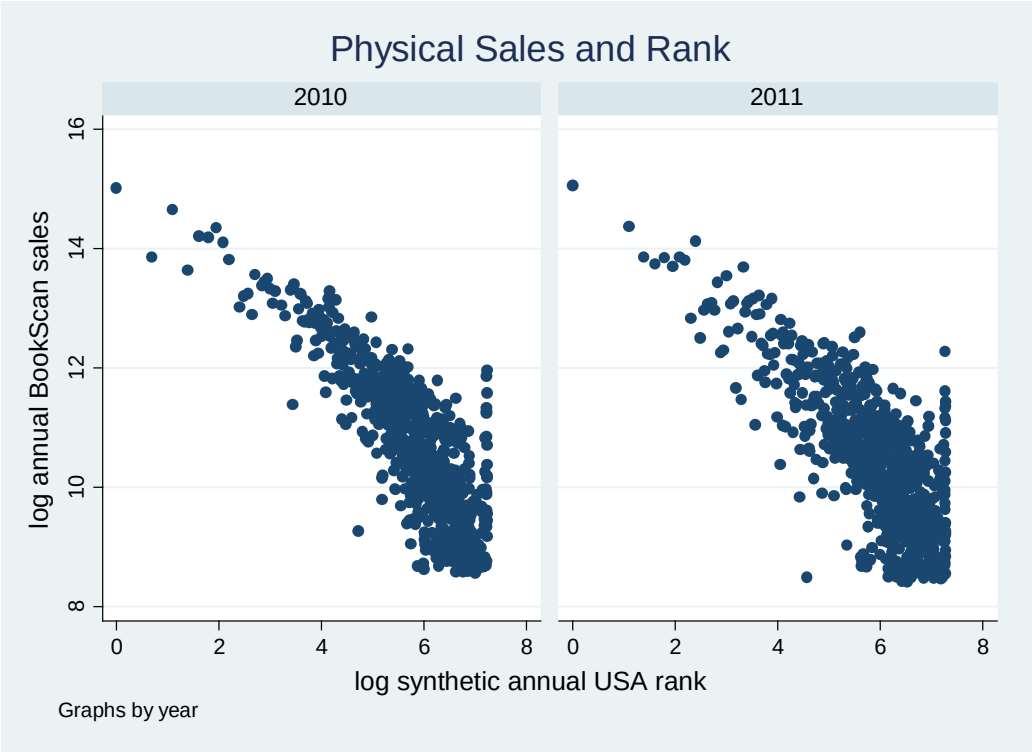


Figure 10

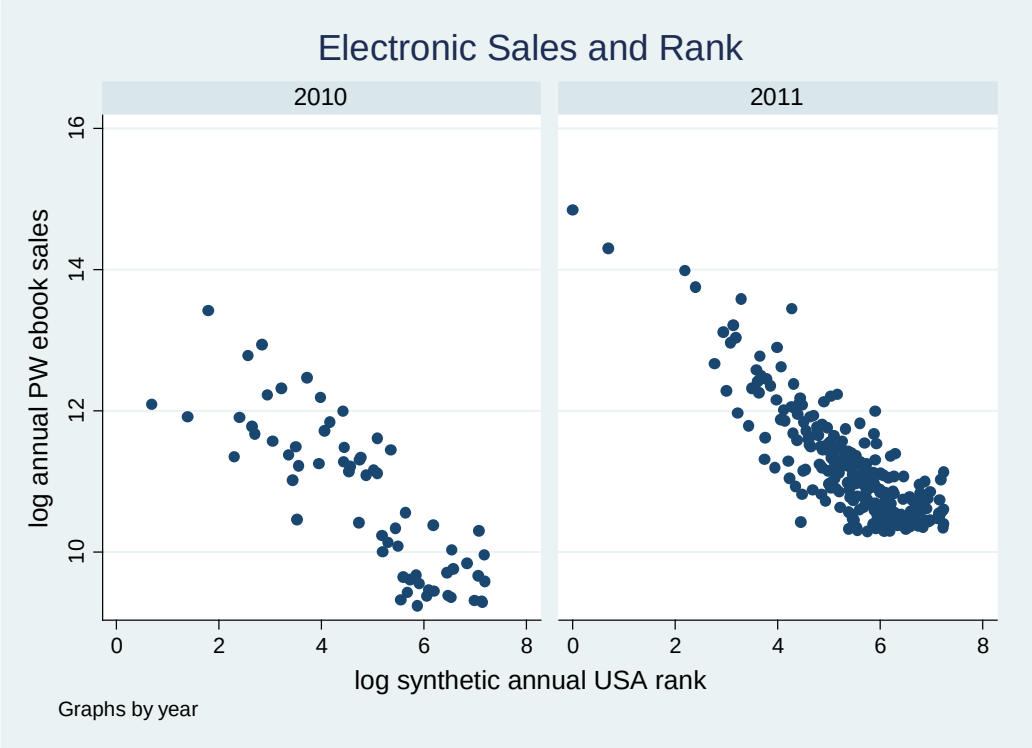


Figure 11

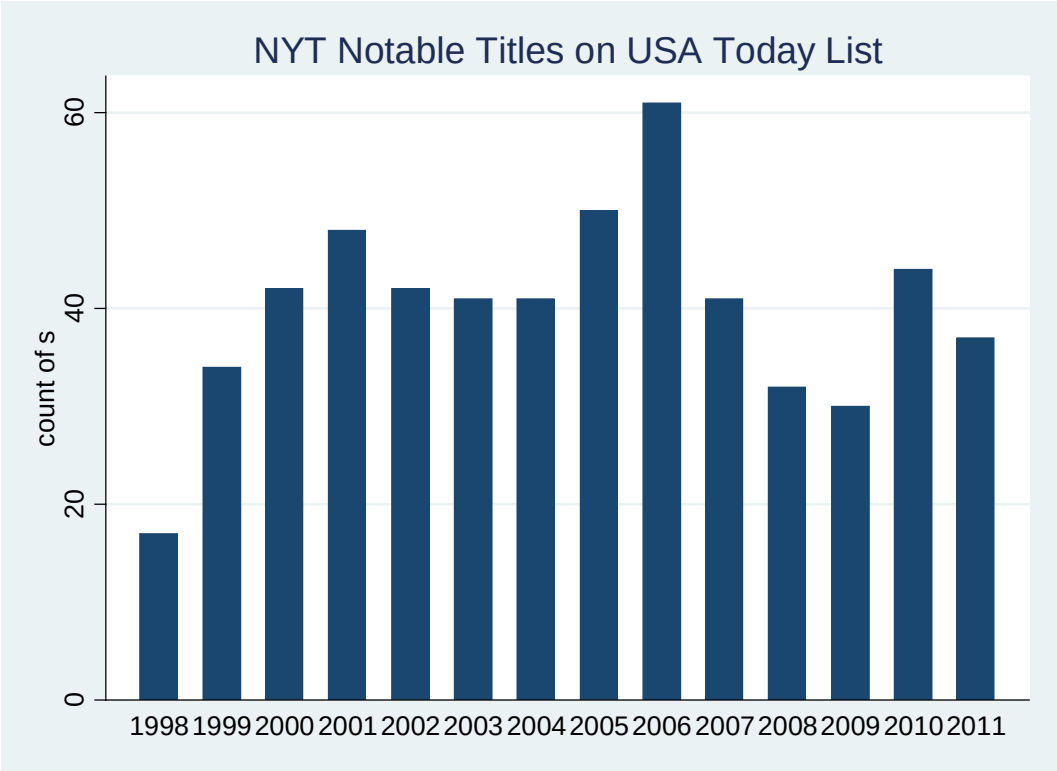


Figure 12

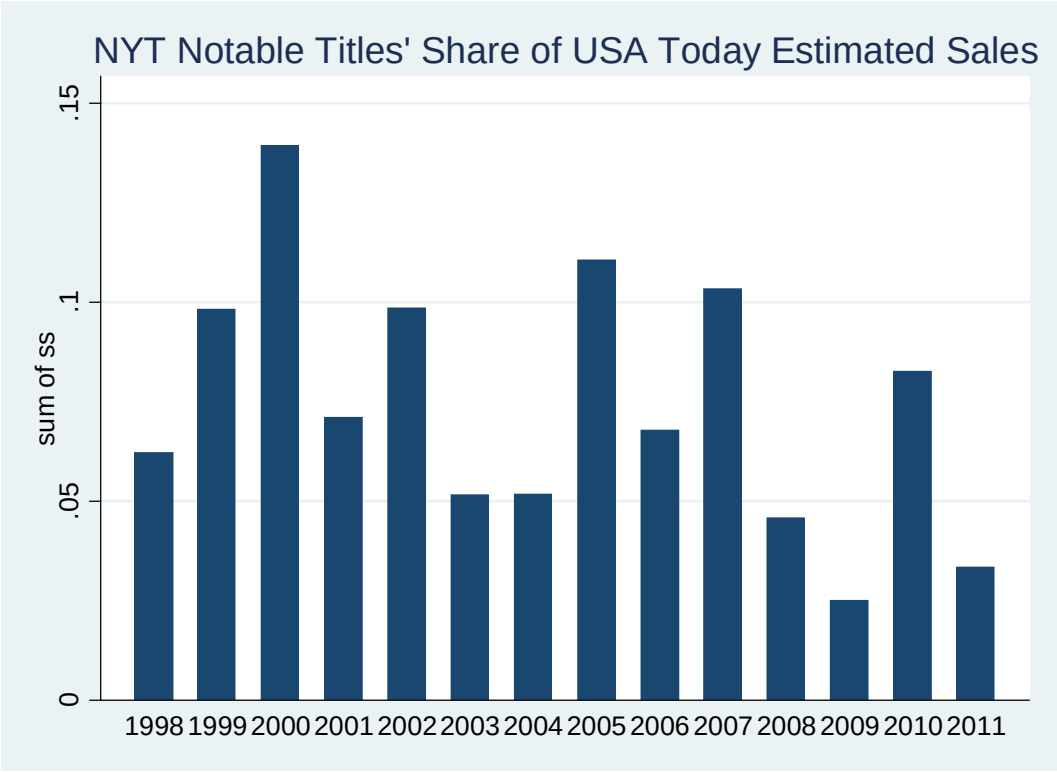


Table 1: Trade Book Sales by Book Format and Type, 2008-2011 (millions of units)

		Format				Type			
	Total	hardcover	mass market	softcover	e-book	Adult Fiction	Adult Nonfiction	Juvenile (Kids & Young Adults)	Religion
2008	2,164.1	570.1	383.4	1,078.0	9.5	593.0	524.8	844.7	201.6
2009	2,211.1	623.3	355.5	1,077.8	35.5	616.2	508.7	909.3	176.9
2010	2,261.3	603.2	319.2	1,099.7	111.9	613.6	497.3	946.6	203.7
2011	2,503.2				388.0				
2012*	2,707.6				531.0				

Source: various media reports on BookStats 2011, 2012. Note that 2012 data are estimated based on August, 2012 year-to-date revenue relative to 2011. The symbol * denotes estimated.
See http://www.publishersweekly.com/pw/by-topic/industry-news/financial-reporting/article/55082-aap-monthly-statshot-august.html .
(See also http://paidcontent.org/2011/08/09/419-new-stats-book-publishing-industry-is-growing-with-e-books-up-over-1000/, http://musingsandmarvels.com/2012/07/18/bookstats-what-happened-in-the-publishing-industry-in-2011/ http://paidcontent.org/2012/07/18/e-books-are-now-the-most-popular-format-for-adult-fiction/).

Table 2: New Book Prices

	total				fiction			
year	hardcover	e-book	trade paperback	mass market	hardcover	e-book	trade paperback	mass market
2008	35.35	59.95	41.12	6.79	30.33	9.10	17.03	6.77
2009	35.78	46.45	41.10	7.02	30.18	8.61	18.16	7.00
2010	35.16	42.92	43.39	7.05	33.22	7.28	18.56	7.01
2011	34.42	27.34	36.98	6.97	29.60	5.31	18.25	6.95

Source: Library and Book Trade Almanac 2012. Prices inflated to 2011 levels using the CPI.

Table 3a: New Titles (Baker & Taylor)

	total titles		fiction titles	
year	physical	electronic	physical	electronic
2008	180,032	35,496	18,638	7,414
2009	178,841	53,731	18,272	13,364
2010	186,344	67,145	17,971	18,043
2011	177,126	111,150	19,760	39,886

Source: Library and Book Almanac 2012.

Table 3b: Total and Self-Published Titles

	Total works		self-published		
	total titles	fiction	total	print	electronic
2002	215,138	25,102			
2003	240,098	24,666			
2004	275,793	38,832			
2005	251,903	34,927			
2006	274,416	42,777	60,875	53,117	7,758
2007	284,370	53,590	74,400	66,459	7,941
2008	289,729	53,058	83,751	75,800	7,951
2009	302,410	48,738	109,019	94,826	14,193
2010	328,259	53,139	149,594	111,551	38,043
2011	347,278	60,075	235,625	148,424	87,201

Sources: Columns 1 and 3 are from Bowker's "New Book Titles and Editions, 2002-2011" (http://www.bowker.com/assets/downloads/products/isbn_output_2002-2011.pdf). The first column ("total") is what Bowker terms "subtotal." It omits "unclassified books, which comprise mostly Reprint/POD houses specializing in public domain works marketed almost exclusively on the web." (http://www.bowker.com/en-US/aboutus/press_room/2012/pr_06052012.shtml)

Table 4: Major Book Review Outlets

Source	2010	2011	target
Booklist	8,457	7,978	librarians
Bookmarks	712	727	general readers, book groups, and librarians
BookPage	828	682	Bookstores and public libraries
Bulletin of the Center for Children's Books	726		school and public librarians
Chicago Tribune Sunday Book Section	500	500	general readers
Choice	6,851	6,833	academic librarians
Horn Book Guide	3,967	4,266	teachers, librarians
Horn Book Magazine	428	482	industry professionals
Kirkus Reviews	4,524	6,197	industry professionals
Library Journal	6,099	6,590	industry professionals
Multicultural Review	378		educators
New York Journal of Books	1,225	1,372	general readers
New York Review of Books	394		general readers
New York Times Sunday Book Review	1,250	1,250	general readers
Publishers Weekly	7,884	7,835	industry professionals
School Library Journal	5,774	6,219	librarians
Washington Post Book World	884	1,136	general readers
TOTAL	50,881	52,067	

Source: Library and Book Trade Almanac 2012, p. 545.

Table 5: Increased Consumer Surplus from Price Reduction

year	price	quantity	Revenue	ΔCS year-to-year	ΔCS rel to 2008	Cumulative ΔCS
2008	14.13	2,164.1	30,569.5			
2009	14.31	2,211.1	31,635.7	-398.1	-398.1	-398.1
2010	13.93	2,261.3	31,501.7	842.8	431.4	33.3
2011	13.19	2,503.2	33,014.6	1,767.2	2,186.1	2,219.4
2012	12.68	2,707.6	34,332.4	1,326.0	3,521.6	5,741.0

Note: all quantities and dollar figures (except price) in millions of 2012 dollars.

Table 6: Imputing E-Book Sales from Overall Ranks

	(1) Log estimated e-book sales 2010	(2) Log estimated e-book sales 2011
log synthetic annual USA Today rank	-0.7041	-0.8058
	(0.0710)**	(0.0505)**
Constant	12.1320	13.8654
	(0.3140)**	(0.2748)**
Observations	1379	1455

Notes: Standard errors in parentheses. * significant at 5%; ** significant at 1% .

Table 7: Sales of Physical and Estimated Sales of Electronic Books on the USA Today Bestseller List

	2009	2010	2011	2012 part year
trade units total (mil)	2,210	2,260	2,500	
trade e-books total (mil)	35.5	111.9	388	
Physical sales on Nielsen weekly top 100	94.9	85.7	72.2	65.5
N	920	954	1036	1060
USA Today Bestseller Sample				
total units (estimate)	98.7	95.2	119	78.5
N	1245	1379	1455	1282
physical	98.7	89.1	78.4	61.6
electronic	0	6.2	40.3	16.9
self-published				
physical	0.022	0.048	1.64	12.9
electronic	0	0.0002	0.47	2.32
self-published share				
physical	0.0%	0.1%	2.1%	20.9%
electronic		0.0%	1.2%	13.7%

Sources: Total and e-book units, BookStats. Physical sales of weekly top 100, Nielsen BookScan. Quantity estimates derived from USA Today ranking are based on annual rank-sales relationships between USA Today overall rankings and a) annualized Nielsen sales data, and b) annual e-book sales for selected titles reported in Publishers Weekly.

Table A1: Originally Self-Published Works and their Sales, 2010-2012 ytd

Author/work	electronic	physical	total
E L James Fifty Shades of Grey	1,051,113	5,466,216	6,517,329
E L James Fifty Shades Darker	343,950	3,216,617	3,560,567
E L James Fifty Shades Freed	248,083	2,841,221	3,089,304
Sylvia Day Bared to You: A Crossfire Nov	141,912	275,870	417,782
Darcie Chan The Mill River Recluse	67,005	264,750	331,755
Chris Culver The Abbey	35,927	154,093	190,019
C.J. Lyons Blind Faith	30,161	126,566	156,728
Amanda Hocking Switched	25,154	89,265	114,419
Jamie McGuire Beautiful Disaster	66,050	43,455	109,505
Amanda Hocking Ascend	21,397	63,097	84,494
Oliver Pötzsch The Hangman's Daughter	18,557	64,762	83,320
Amanda Hocking Torn	19,164	63,755	82,918
Michael Prescott Last Breath	18,321	61,325	79,646
Colleen Hoover Slammed	20,551	55,748	76,299
Michael Prescott Stealing Faces	15,007	57,709	72,715
Rick Murcer Caribbean Moon	14,823	56,915	71,739
Michael Prescott Blind Pursuit	13,567	51,518	65,085
Catherine Bybee Wife by Wednesday	15,871	45,936	61,807
Colleen Hoover Point of Retreat	16,374	43,441	59,815
Lara Adrian Deeper Than Midnight	45,018	12,843	57,861
Barbara Freethy Summer Secrets	11,976	44,773	56,749
Barbara Freethy The Sweetest Thing	15,389	40,582	55,972
Tammara Webber Easy	14,357	37,605	51,963
Tracey Garvis-Graves On the Island	41,013	10,713	51,726
P.J. Alderman A Killing Tide	10,634	39,169	49,803
Stephanie McAfee Diary of a Mad Fat Girl	10,119	37,041	47,160
Nancy C. Johnson Her Last Letter	9,366	33,958	43,324
Barbara Freethy Silent Run	9,239	33,438	42,677
R.L. Mathewson Playing for Keeps	11,053	28,219	39,272
Abbi Glines While It Lasts	10,809	27,537	38,346
Kelli Maine Taken	6,654	31,414	38,068
Barbara Freethy Love Will Find A Way	7,719	27,314	35,033
Bella Andre If You Were Mine: The Sulliv	9,894	24,988	34,882
Amanda Hocking My Blood Approves	7,677	27,148	34,825
Barbara Freethy Golden Lies	9,870	24,921	34,790
Barbara Freethy Don't Say A Word	7,663	27,093	34,757
Barbara Freethy Daniel's Gift	7,609	26,876	34,485
Lynda Chance Sarah's Surrender	9,703	24,460	34,163
Chris Culver Just Run	7,463	26,298	33,761
Barbara Freethy All She Ever Wanted	8,413	24,552	32,965

Kathleen Long Chasing Rainbows	9,366	23,530	32,896
Sydney Landon Weekends Required	8,747	21,828	30,576
Bella Andre Can't Help Falling In Love:	8,497	21,142	29,639
Michael Prescott Mortal Pursuit	6,575	22,805	29,379
Lyla Sinclair Training Tessa	8,377	20,817	29,194
Gemma Halliday High Heels Mysteries Boxe	8,360	20,771	29,131
Tina Reber Love Unrehearsed	8,327	20,680	29,007
Courtney Milan Unlocked	6,158	21,184	27,342
Molly McAdams Taking Chances	7,847	19,374	27,221
Bella Andre Let Me Be The One: The Sulli	7,555	18,584	26,139
Lara Adrian Darker After Midnight	7,399	18,163	25,562
Susan Ee Angelfall (Penryn & the End of	7,336	17,993	25,329
Evan Katy January Kills Me	5,591	19,003	24,594
Amanda Hocking Fate	5,544	18,823	24,367
Barbara Freethy Silent Fall	5,517	18,722	24,239
Victorine E. Lieske Not What She Seems	5,504	18,672	24,176
Kelly Favor For His Pleasure	6,829	16,635	23,464
Rick Murcer Deceitful Moon	5,303	17,906	23,209
C.J. Lyons Borrowed Time	5,238	17,658	22,895
Bella Andre The Look of Love: The Sulliv	6,594	16,008	22,602
Michael Prescott Riptide	5,025	16,853	21,878
Amanda Hocking Flutter	4,983	16,694	21,676
Heather Killough-Wal The Spell	4,967	16,635	21,602
Lara Adrian Taken by Midnight: A Midnigh	96	21,493	21,589
J.R. Rain Samantha Moon	6,261	15,121	21,382
H.P. Mallory The Witch is Back	6,166	14,870	21,037
Lara Adrian Lord of Vengeance	6,099	14,694	20,793
Barbara Freethy Just the Way You Are	6,083	14,650	20,733
Rebecca Donovan Reason to Breathe	5,994	14,416	20,410
Heather Killough-Wal The Heat	4,707	15,659	20,367
Jennifer Ashley Hard Mated	5,978	14,374	20,353
Ruth Cardello Bedding the Billionaire	5,788	13,873	19,662
R.L. Mathewson Perfection: A Neighbor Fr	5,653	13,517	19,169
Mimi Jean Pamfiloff Accidentally Married	5,524	13,179	18,703
Robyn Carr Chelynne	4,348	14,323	18,671
Barbara Freethy Some Kind of Wonderful	4,175	13,682	17,857
John Locke Saving Rachel	4,059	13,254	17,313
Hugh Howey Wool Omnibus Edition (Wool 1	5,085	12,033	17,118
Maya Banks Softly at Sunrise	4,993	11,796	16,789
Kelly Favor His Every Desire	4,957	11,701	16,658
Sara Fawkes Anything He Wants	4,921	11,608	16,529
Lara Adrian A Taste of Midnight: A Midni	3,859	12,521	16,380
Sarah Burleton Why Me?	4,599	10,779	15,378

J.R. Rain Vampire Dawn	4,480	10,473	14,953
Kelly Favor His Every Touch	4,360	10,165	14,525
Jessica Park Flat-Out Love	4,322	10,066	14,388
Barbara Freethy Taken	3,375	10,771	14,147
Kelly Favor For His Forever	4,179	9,701	13,879
Jennifer Ashley Primal Bonds	3,290	10,466	13,756
Lara Adrian Shades of Midnight	73	13,319	13,392
Jennifer Ashley Pride Mates	38	13,273	13,311
Jennifer Ashley The Duke's Perfect Wife	4,010	9,272	13,282
Debbi Mack Identity Crisis	3,150	9,967	13,117
Heather Killough-Wal The Strip	3,119	9,855	12,974
H.P. Mallory Witchful Thinking	3,916	9,033	12,949
Michael Prescott Mortal Faults	3,891	8,972	12,863
Amanda Hocking Wisdom	3,039	9,570	12,609
Kallypso Masters Nobody's Perfect	3,750	8,614	12,364
Nicole Williams Clash	3,703	8,497	12,200
Rachel Astor Gamble on Engagement	3,698	8,484	12,182
Sara Fawkes Anything He Wants 4: Collate	3,653	8,370	12,023
Jennifer Ashley Wild Cat	3,630	8,311	11,941
Abbi Glines Ceaseless	3,594	8,222	11,816
Kathleen Brooks Rising Storm	3,579	8,184	11,763
Sara Fawkes Anything He Wants 2: All's F	3,537	8,079	11,616
Kelly Favor For His Honor	3,535	8,073	11,607
Debra Holland Wild Montana Sky	3,482	7,941	11,423
Abbi Glines Because of Low	3,456	7,877	11,333
Nicole Williams Crash	3,447	7,854	11,301
Terri Giuliano Long In Leah’s Wake	3,438	7,831	11,269
TOTAL	2,801,352	14,622,461	17,423,813

References

Bittlingmayer, Georg

Blackburn, David. “On-line Piracy and Recorded Music Sales.” Unpublished Manuscript, Harvard University, December 2004.

Brynjolfsson, Erik, Michael D. Smith, and Yu (Jeffrey) Hu, "Consumer Surplus in the Digital Economy: Estimating the Value of Increased Product Variety at Online Booksellers" *Management Science*, Vol. 49 (2003).

Chevalier, Judith, and Austan Goolsbee, “Measuring Prices and Price Competition Online: Amazon vs. Barnes and Noble,” *Quantitative Marketing and Economics I,* 2 (2003), 203–222.

Caves, Richard E. Creative Industries: Contracts between Art and Commerce. Harvard University Press: Cambridge, MA. 2000.

Greco, Albert N. The Book Publishing Industry, 2ed.New York, Lawrence Erlbaum Associates. 2005.

Heald

Liebowitz, Stan J. "File Sharing: Creative Destruction or Just Plain Destruction?" Journal of Law and Economics, vol. 49, no. 1, April 2006, pp. 1-28.

Oberholzer-Gee, Felix and Koleman Strumpf. “The Effect of File Sharing on Record Sales: An Empirical Analysis.” Journal of Political Economy. Volume 115, Issue 1, Page 1–42, Feb 2007

Rainie, Lee, Kathryn Zickhur, Kristen Purcell, Mary Madden, and Joanna Brenner. 2012. “The Rise of e-Reading.” Pew Research Center, Washington.

Rainie, Lee. 2012. “Tablet and E-book reader Ownership Nearly Double Over the Holiday Gift-Giving Period.” Pew Research Center, Washington.

Rob, Rafael and Joel Waldfogel. “Piracy on the High C’s: Music Downloading, Sales Displacement, and Social Welfare in a Sample of College Students.” *Journal of Law & Economics*. Volume 49, Issue 1, Page 29-62, Apr 2006.

Smith

Sorenson

Tervio, Marko. “Superstars and Mediocrities: Market Failure in the Discovery of Talent.” Review of Economic Studies 72 (2): 829-850, 2009.

Vogel, Harold, Entertainment Industry Economics, 7th edition. Cambridge, Cambridge University Press, 2007.

Waldfogel, Joel. “Copyright Protection, Technological Change, and the Quality of New Products: Evidence from Recorded Music since Napster.” Forthcoming. *Journal of Law & Economics*.

Waldfogel, Joel. “And the Bands Played On.” Forthcoming. The Economics of Digitization: An Agenda. Shane Greenstein, Avi Goldfarb, and Catherine Tucker, eds.

Zentner, Alejandro, “Measuring the Effect of File Sharing on Music Purchases.” The Journal of Law and Economics. Volume 49, Issue 1, Page 63–90, Apr 2006.