

This PDF is a selection from a published volume from the National Bureau of Economic Research

Volume Title: Economic Analysis of the Digital Economy

Volume Author/Editor: Avi Goldfarb, Shane M. Greenstein, and Catherine E. Tucker, editors

Volume Publisher: University of Chicago Press

Volume ISBN: 0-226-20684-X; 978-0-226-20684-4

Volume URL: <http://www.nber.org/books/gree13-1>

Conference Date: June 6–7, 2013

Publication Date: April 2015

Chapter Title: Searching for Physical and Digital Media: The Evolution of Platforms for Finding Books

Chapter Author(s): Michael R. Baye, Babur De los Santos, Matthijs R. Wildenbeest

Chapter URL: <http://www.nber.org/chapters/c12989>

Chapter pages in book: (p. 137 – 165)

---

# Searching for Physical and Digital Media

## The Evolution of Platforms for Finding Books

Michael R. Baye, Babur De los Santos,  
and Matthijs R. Wildenbeest

---

### 5.1 Introduction

The Internet has had a profound impact on the way people shop; online sales in the United States were \$213 billion in 2012, accounting for 8 percent of all retail sales.<sup>1</sup> The Internet and digitization have had an even greater impact on sales of media products, thanks in part to changes in the way people search for, acquire, and consume these products. Online sales of books accounted for about 44 percent of all US book sales in 2012, with e-books accounting for 11 percent of total book sales.<sup>2</sup> Digital sales of music accounted for over half of the industry's 2012 total sales.<sup>3</sup>

This chapter focuses on how the Internet and digitization are changing the way consumers obtain information about media products such as books, music, and movies. Less than two decades ago, a consumer looking for a

Michael R. Baye is the Bert Elwert Professor of Business and a professor of business economics and public policy at the Kelley School of Business, Indiana University. Babur De los Santos is assistant professor of business economics and public policy at the Kelley School of Business, Indiana University. Matthijs R. Wildenbeest is associate professor of business economics and public policy at the Kelley School of Business, Indiana University.

We thank two anonymous referees, as well as Avi Goldfarb, Shane Greenstein, Marc Rysman, Catherine Tucker, Hal Varian, and seminar participants at the NBER Economics of Digitization conference for valuable comments. We also thank Susan Kayser, I. K. Kim, Joowon Kim, Yoojin Lee, and Sarah Zeng for research assistance. Funding for the data and research assistance related to this research was made possible by a grant from Google to Indiana University. The views expressed in this chapter are those of the authors and do not necessarily reflect the views of Indiana University or Google. For acknowledgments, sources of research support, and disclosure of the authors' material financial relationships, if any, please see <http://www.nber.org/chapters/c12989.ack>.

1. Figures from Forrester Research, March 2013.
2. Figures from *Publishers Weekly*, May 2013.
3. Figures from Nielsen, January 2013.

book, CD, or DVD had to spend considerable time searching among sellers of these physical products. Given search technologies at the time (e.g., using a telephone to call different sellers or traveling from one physical store to another), the typical consumer limited purchase decisions to retailers proximate to his or her domicile. The Internet and digitization have reduced consumers' costs of obtaining information, thus making it easier for consumers to identify and purchase media that best match their preferences.

Leading media platforms (e.g., Netflix, Amazon, and iTunes) now provide users with recommendations for specific media titles based on sophisticated algorithms that account for a given user's preferences as well as the ratings provided by other users. Consumers also receive recommendations through social networks such as Facebook; "friends" can indicate whether they "liked" a particular song, movie, or book. Many of these recommendation systems operate in real-time and use purchase patterns of consumers with similar profiles to identify products consumers would be otherwise unaware of.

Additionally, advances in search technologies have made it easier for consumers to directly access information that is relevant for their purchase decisions. Consumers can easily search for product information by using a general search engine (e.g., Google), a price comparison site (e.g., Book-Finder), as well as by navigating to a specific retailer's site and conducting a search on its platform (e.g., Amazon and iTunes). These changes in search technologies allow consumers to find a greater breadth of products and make the long tail more accessible; it is now easier to find rare and obscure books, music, and movies. And thanks to digitization, consumers can now view samples of book pages, listen to sample music tracks, and watch scenes from movies through a plethora of devices connected to the Internet. In short, finding the "right" product now takes less effort, and it is easier to compare the prices different sellers charge for that product.

This chapter is motivated by more recent changes in the way consumers search for media products. Until recently, most online product searches for media were conducted using browser-based platforms. Increasingly, however, searches for digital media take place on closed platforms that sell music and video content (e.g., iTunes) or books (e.g., the Amazon Kindle and Barnes & Noble Nook). This shift is more mature in the music industry; digital content accounted for 37 percent of music sales in 2012—more than three times the share of digital books among total book sales.

Because digital books are still in their infancy, the remainder of this chapter focuses on the impact of the Internet and digitization on the book industry. Our goal is to provide readers with a glimpse of how searches for books are likely to evolve over the next decade and to highlight some of the challenges and questions that researchers working in this area are likely to encounter in the years to come. Nonetheless, many of the issues that we discuss are relevant for media such as music and videos, as well as nonmedia

products that consumers may search for and purchase online. We highlight a number of data limitations and technical challenges associated with identifying product and product-related searches on browser-based platforms (e.g., search engines, price comparison sites, and retail websites), as well as closed systems (e.g., applications, or apps, on smartphones and tablets such as the iPad, and platforms such as the Kindle and Nook). Finally, we provide some preliminary evidence that searchers are shifting toward new and evolving search platforms and away from browser-based platforms (such as search engines) to find books.

As alluded to above, the book industry is a useful one to examine because it has already felt significant changes as a result of the Internet and is likely to experience additional changes over the next decade as a result of digitization. Even though e-books are still in their infancy, this technology allows consumers to locate, purchase, and read books virtually anywhere. The challenge for researchers over the next decade is that this activity increasingly occurs on closed systems where search activity is not readily observable by those “outside” of the platform.

Our chapter is related to a vast literature on consumer search behavior that began with Stigler’s (1961) seminal article on the economics of information. Subsequent literature has focused on explaining price dispersion as an equilibrium phenomenon—particularly in online markets where informational frictions are very modest.<sup>4</sup> The more recent literature is decidedly empirical and examines how industry characteristics, such as search costs, impact levels of price dispersion. For instance, at the beginning of the millennium a large number of studies examined the effects of the Internet on prices and price dispersion.<sup>5</sup> Several of these studies focused specifically on the online book industry. Clay, Krishnan, and Wolff (2001) examined prices at online bookstores and found that heightened competition among sellers led to lower prices and less price dispersion. Moreover, they established heterogeneity in firm-level behavior. Brynjolfsson, Hu, and Smith (2003) quantify the economic impact of increased variety at online bookstores (and the ability of consumers to identify books in the long tail). Their estimates indicate that online markets for books increased consumer welfare by as much as \$1 billion for the year 2000. Chevalier and Goolsbee (2003) show how to use sales rank data to obtain elasticity estimates for Amazon and Barnes & Noble, and find that demand is much more price elastic for Barnes & Noble than Amazon. Chevalier and Mayzlin (2006) study the effect of consumer reviews on sales at amazon.com and barnesandnoble.com. Hong and Shum (2006) develop techniques to estimate search costs using only price data and apply their methods to books sold online. Forman, Ghose,

4. See Baye, Morgan, and Scholten (2006) for a survey.

5. See Baye, De los Santos, and Wildenbeest (2013) for an overview of the evolution of product search in both offline and online markets.

and Goldfarb (2009) provide powerful evidence that online and offline markets for books are linked: when a physical book store opens, consumers near the area purchase fewer books online. De los Santos (2012) and De los Santos, Hortaçsu, and Wildenbeest (2012) use data on browsing behavior and book purchases to estimate search costs.

The remainder of this chapter is organized as follows. The next section provides a brief overview of the book industry and highlights some of the more important changes that have occurred over the past twenty years. Section 5.3 provides a data-driven look at how consumers use different browser-based platforms (such as websites of booksellers, price comparison sites, and search engines) to locate books and booksellers. Section 5.4 investigates how search intensity changes throughout the life cycle of a book and provides preliminary evidence that searches for book titles on general search engines are beginning to decline. Specifically, consumers appear to be shifting away from general search engines and are increasingly using retailer sites and closed systems to find books. Section 5.5 provides a look at differences in the prices of digital and physical books. It also examines how these prices evolved before, during, and after the Antitrust Division of the Department of Justice filed suit against Apple and book publishers regarding their use of agency contracts. We conclude in section 5.6 by briefly describing several additional agenda items for research programs.

## **5.2 Overview of the Book Industry and Description of Data Sets**

The book industry has seen many changes throughout the past twenty years. The most significant change, to date, has been the move from selling physical books in traditional brick-and-mortar bookstores to selling physical books online. Books were one of the first products that were successfully sold online. One of the earliest and largest online retailers, Amazon, started as an online bookstore in July 1995. Barnesandnoble.com was launched in May 1997. Many startups followed, and existing brick-and-mortar bookstores began creating an online presence. Most of these online bookstores (retailers such as 1bookstreet, A1books, and Fatbrain) no longer exist; many went bankrupt during the collapse of the Internet bubble shortly after the new millennium began. Borders, which was one of the largest brick-and-mortar bookstores in the United States, went bankrupt in 2011. Today, the retail book industry consists of Amazon, Barnes & Noble, and smaller bookstores as well as mass merchants (such as Walmart), which typically have thinner offerings of titles.

A more recent change in the book industry is the shift toward electronic books, or e-books. Although e-books have been around since at least 1971, when Michael Hart made the Declaration of Independence electronically available as part of Project Gutenberg, the release of the Sony LIBRIé e-book reader in 2004 in Japan (the United States followed in 2006 with

**Table 5.1** US print and e-book sales, 2006–2009

Year	Trade print		E-books		Total
	(\$ in millions)	(%)	(\$ in millions)	(%)	(\$ in millions)
2006	11,123	99.5	54	0.5	11,177
2007	11,384	99.4	67	0.6	11,451
2008	10,831	99.0	113	1.0	10,944
2009	10,711	97.2	313	2.8	11,024

Source: Mintel/American Association of Publishers. E-book sales exclude educational and professional books.

the release of the Sony PRS-500) marked the beginning of the recent surge in popularity of e-books.<sup>6</sup> The Sony LIBRIé was the first commercially available device to use E Ink technology, which dramatically improved the reading experience and required less battery power than existing e-readers at the time. Amazon released its E Ink technology-based Kindle in November 2007. Barnes & Noble followed in November 2009 with the release of the Nook. At the end of June 2012, Amazon led the market for devices used for reading e-books, with an overall market share of 55 percent (with its Kindle and Kindle Fire), followed by Nook (14 percent). Today, Sony's e-book readers have 1 percent of the market.<sup>7</sup> The definition of an e-reader is increasingly blurred by the presence of devices such as tablets, smartphones, desktops, and laptops with applications for reading e-books. Indeed, after having seen dramatic growth since 2006, the sales of traditional e-readers are expected to decline in 2012, with consumers shifting toward tablets.<sup>8</sup>

The sales of e-books have grown alongside the increasing popularity of e-readers. By July 2010, Amazon was selling more Kindle books than hard-cover books, and since April 2011 it is selling more Kindle books than hard-cover and paperback books combined (excluding free e-books).<sup>9</sup> However, as shown in table 5.1, e-books accounted for a very modest share of overall industry sales between 2006 and 2009. Between 2009 and 2012, however, sales of e-books grew from 2.8 percent to 11 percent of total industry sales.

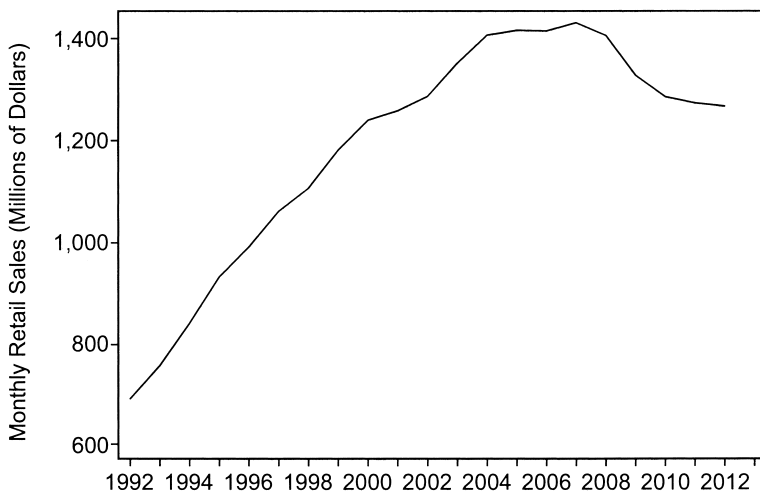
Most e-readers support the open ePub format (Amazon's Kindle, which supports the proprietary AZW format, is a notable exception). In theory, this means consumers can buy an e-book at one of the bookstores that supports the ePub format and read it on a different e-reader. However, bookstores use Digital Rights Management (DRM) to protect the content of the e-books,

6. See Manley and Holley (2012) for an overview of the history of e-books.

7. Estimates from Bowker Market Research.

8. See [http://news.cnet.com/8301-1035\\_3-57558710-94/rip-e-book-readers-rise-of-tablets-drives-e-reader-drop/](http://news.cnet.com/8301-1035_3-57558710-94/rip-e-book-readers-rise-of-tablets-drives-e-reader-drop/).

9. See <http://phx.corporate-ir.net/phoenix.zhtml?c=176060&p=irol-newsArticle&ID=1565581>.



**Fig. 5.1 Monthly retail sales at brick-and-mortar bookstores**

*Source:* Monthly Retail Trade Survey for NAICS 451211 (bookstores), US Census Bureau.

and different bookstores do not necessarily use the same DRM standard. Even though there are ways to get around this (for instance, by stripping the DRM), for most consumers this is difficult to do. This means that in practice the e-book market has many characteristics of a two-sided market, with publishers on one side of the market, readers on the other side, and bookstores and e-readers as intermediaries.<sup>10</sup>

The rise of e-books is likely to significantly impact the way consumers search for books. Buyers of e-books are less likely to use web browsers to find and purchase books, as it is often more convenient to search within the device that is used to read e-books. Moreover, the lack of interchangeability due to DRM makes it less beneficial for consumers to compare prices across bookstores or platforms, resulting in fewer retailers being visited. Unfortunately, we do not observe consumer search patterns within most of the (closed) platforms, and future researchers are likely to increasingly experience this problem not only for searches for books, but searches for any product that take place through a closed device or mobile device app. Our challenge here is to attempt to tease out—from available data on browser-based searches—evidence of shifts toward searches on platforms or devices for which search data are unavailable. Thus, the remainder of this chapter focuses exclusively on browser-based searches.

These developments—the move toward selling books online as well as the transition to e-books—have significantly impacted the sales of books at traditional brick-and-mortar stores. Figure 5.1 shows that average

10. See Rysman (2009) for a discussion of the economics of two-sided markets.

**Table 5.2** Revenue of leading bookstores, 2005–2009

	Revenue (\$ in millions)					Change (%)
	2005	2006	2007	2008	2009	
<i>Bookstore chains</i>						
Barnes & Noble	5,103	5,261	5,411	5,122	5,596	9.7
Borders	4,031	4,063	3,775	3,242	2,791	–30.8
Books-A-Million	504	521	536	515	509	1.0
<i>Online bookstores</i>						
Amazon.com	3,046	3,582	4,630	5,350	5,964	95.8
Barnesandnoble.com	440	433	477	466	525	19.4
Borders.com	—	—	—	46	60	—
Booksamillion.com	28	26	27	25	24	–13.8

*Source:* Simba information, from company reports. Figures for Amazon.com are for US books, music, and video/DVD segment only. Until 2008 Borders.com was powered by Amazon.com.

monthly sales at brick-and-mortar bookstores roughly doubled between 1992 and 2007. Since 2007, however, sales at traditional bookstores have been in decline. This decline in the sales of physical books coincides with two events, making it difficult to separately identify the cause of the decline. One event was a recession, which according to the National Bureau of Economic Research (NBER) started in December 2007 and ended in June 2009. The other event was Amazon’s introduction of the Kindle e-book reader in November 2007. While we view the cause of this decline as an interesting and open question, the data in table 5.2 suggest that the post-2007 decline in sales at traditional bookstores may, in part, be due to a shift toward e-books and the online channel more generally.

The analysis in the remainder of this chapter relies on four different data sets. The first data set is assembled using data from comScore Search Planner and contains information on the most popular search terms that are used at the main general search engines to reach specific websites for the period between February 2010 and February 2013.<sup>11</sup> Search Planner uses the comScore panel, which contains all online browsing activity of around two million US users. Our second data set, comScore qSearch, also uses the comScore panel and contains monthly data (between January 2011 and February 2013) on search volume and search intensity at over 200 online properties.<sup>12</sup> The third data set uses relative search volume data from Google Trends for the period 2004–2013. Our final data set contains prices for printed and

11. For a more detailed description of these data, see Baye, De los Santos, and Wildenbeest (2012).

12. For a more detailed description of these data, see Baye, De los Santos, and Wildenbeest (2013).



digital books for a large number of titles sold at the major bookstores for 2012. These data were obtained using a scraper written in Java.<sup>13</sup>

With this brief background, we now turn to search.

### 5.3 Searching for Books and Booksellers

There are numerous ways consumers can use the Internet to search for books or a bookseller. Consumers can use a general search engine such as Google, Yahoo, or Bing to search for a specific book title, author name, or other book-related search term. When using a search engine, search results typically appear as links that can be clicked, which redirect the user to a specific (often external) website. Most of the general search engines display both paid and nonpaid (organic) results; paid results typically appear on top of the search results page as well as on the side.

An alternative way to search is by using a price comparison site to identify sellers of a particular book. Price comparison sites are intermediaries that provide consumers with price information as well as other product information.<sup>14</sup> Most price comparison sites are free for consumers to use, while retailers pay fees. Typically, these fees are paid each time a searcher clicks a link at the price comparison site that directs the user to the retailer's site. Most comparison sites allow users to either browse or use a search box to search. Consumers searching for books can choose among many different comparison sites, ranging from general price comparison sites such as Nextag and Bing Shopper to book-specific comparison sites such as BooksPrice or BookFinder. The more book-oriented price comparison sites typically allow users to search by ISBN as well as title or author name.

A third option is to go directly to a bookseller's (or other retailer's) website and conduct a search within its site. Most online bookstores allow consumers to search by title, author name, or ISBN. Online bookstores such as Amazon and Barnes & Noble also provide links to bestselling books on their web pages, allowing consumers to explore product offerings without having to use the search box. In addition, suggestions based on previous purchases or items previously viewed are displayed to help consumers find the right book. Some online booksellers also allow third-party sellers to sell books on their websites (e.g., Amazon Marketplace and Barnes & Noble Marketplace), which allows consumers to do a price comparison without having to leave the site.

As we noted earlier, an increasingly attractive fourth option—especially for those searching for e-books—is to conduct a search on a closed device

13. For a more detailed description of these data, see De los Santos and Wildenbeest (2014).

14. See Moraga-González and Wildenbeest (2012) for a discussion on comparison sites and their implications for price competition.

such as Amazon's Kindle or the Barnes & Noble Nook. Recall that these search data are not readily available to researchers.

The remainder of this section examines the relative importance of the first three types of book searches and highlights a number of obstacles that make it difficult to precisely measure the share of book searches that are conducted on different search platforms. Later, we will attempt to use the data that are available to examine whether recent trends in general search engine use suggest that consumers are indeed shifting to closed search platforms, for which detailed search data are presently unavailable.

### 5.3.1 Book-Related Searches on General Search Engines

We first turn to the issue of measuring the number of book-related searches on the main general search engines (Google, Bing, Yahoo, and Ask). For this we use comScore's Search Planner data, which is based on the browsing activity of two million US users that are part of the comScore panel. As part of this data we observe the most popular search terms used at the main general search engines to reach a specific site. To illustrate the nature of the data, table 5.3 shows the top twenty-five search terms on Google that resulted in organic traffic to Barnes & Noble.<sup>15</sup> The number one search phrase is "barnes and noble," which generated 376,678 clicks to this site. Notice that several other top twenty-five search terms in table 5.3 are variations or misspellings of the site name or URL (e.g., "barnes and nobles," "barnes," "bn," and "www.bn.com"). This suggests that searchers may frequently use general search engines as a shortcut for typing in the URL of the site they wish to visit, and then conduct actual searches for books at the site. Also notice that the list of top twenty-five search terms includes competitors (e.g., "amazon"). Farther down the list we begin seeing terms related to specific books (e.g., the 20th ranked search term is "the vow book").

Table 5.3 also illustrates a weakness of the Search Planner data: for privacy and other reasons, comScore sometimes reports "\*\*\*\*" and "name name" rather than the actual search phrase. This is the case for the fourth and sixth ranked search terms in table 5.3, and as a consequence, one cannot precisely determine the nature of these searches. For purposes of the analysis that follows, these terms are ignored. This is not ideal, since these unknown (and hence excluded) searches may refer to an author or ISBN number.

In order to get a better picture of the nature of search terms leading users from search engines to bookstores, we examined the entire list of search terms resulting in traffic to Amazon and Barnes & Noble (2,053 and 180

15. In September 2011, Barnes & Noble acquired Borders's intellectual property, including Borders's domain name. Borders.com now redirects to barnesandnoble.com, explaining the relatively large number of Borders's related search terms in table 5.3.

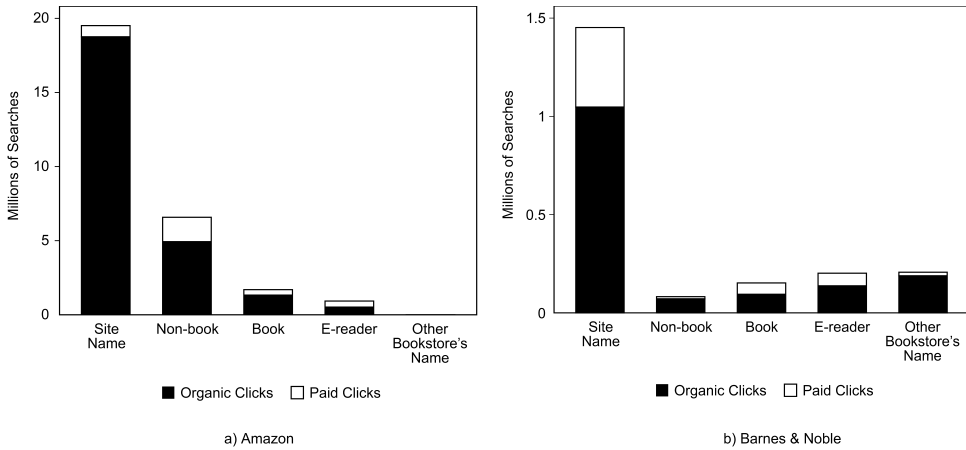
**Table 5.3** Top twenty-five search terms on Google leading users to Barnes & Noble

Rank	Search phrase	No. organic clicks
1	barnes and noble	376,678
2	borders bookstore	82,689
3	borders book store	52,006
4	***	27,699
5	barnes and noble locations	21,666
6	name name	20,675
7	barnes and nobles	19,800
8	amazon	19,748
9	barnesandnoble	17,785
10	nook	13,483
11	facebook	10,680
12	barnes & noble	9,900
13	nook tablet	9,775
14	barnes	8,623
15	borders.com	7,465
16	google	7,163
17	bn.com	7,118
18	books	6,923
19	borders	6,382
20	the vow book	6,199
21	name fire	5,143
22	bn	4,819
23	nook color update	4,564
24	gmail	4,383
25	www.bn.com	4,179

*Source:* comScore Search Planner data from February 2012. Search phrases are ranked by the total number of organic clicks on Google.

search terms, respectively), and grouped them into five categories: (1) site name, which includes the name or URL (including misspellings) of the site; (2) nonbook, which includes products other than books; (3) book, which includes physical as well as e-books; (4) e-reader, which includes searches for e-reader-related hardware or software; and (5) other bookstore, which includes names and URLs of other bookstores (including misspellings). The results are displayed in figure 5.2, panels (a) and (b), and provide a breakdown of the observed search terms consumers entered at the main search engines that resulted in clicks to Amazon and Barnes & Noble, respectively.

Several features of figure 5.2 are worth noting. First, the overwhelming majority of searches at the main search engines leading users to Amazon or Barnes & Noble were navigational searches, that is, search terms that only include the name or URL of the retailer. This means that most of the traffic Amazon and Barnes & Noble receive from the main search engines is from consumers using search engines instead of the address bar to navigate. Pre-



**Fig. 5.2 Search terms in main search engines leading users to Amazon and Barnes & Noble**

*Source:* comScore Search Planner, February 2012.

sumably, users conduct specific searches for books once they land on these booksellers' sites.

Second, notice in figure 5.2, panel (a), that searches that are not book related generate more traffic to Amazon than book-related searches. This is consistent with the fact that Amazon is now a mass merchant that sells many products other than books. Figure 5.2, panel (b), shows that book-related search terms generate far more traffic to Barnes & Noble than search terms that are not book related. This is consistent with the fact that Barnes & Noble specializes in books. Finally, figure 5.2 reveals that searches for e-readers and other bookstores are more important for Barnes & Noble than for Amazon in generating traffic to its site. The key takeaway from figure 5.2 is that navigational searches account for much of the traffic that Amazon and Barnes & Noble obtain from general search engines.

We conclude this section with a caveat that is related to another drawback of the Search Planner data used in our analysis: comScore reports search terms leading users to a particular site only if the number of clicks exceeds a certain threshold. Thus, while we do observe the most popular search terms generating traffic to a specific bookstore, we do not observe the more obscure searches taking place on general search engines. For instance, for Amazon we only observe the search terms corresponding to 24 percent of total click volume. Brynjolfsson, Hu, and Simester (2011) have shown that Internet markets have increased the importance of niche markets in terms of sales. This long-tail phenomenon suggests that a portion of the unobserved searches may include searches for book titles and other more obscure book-

**Table 5.4** Percentage of incoming visits to bookstores from search engines and other platforms

	Amazon (%)	Barnes & Noble (%)	eCampus (%)	Powell's (%)	Books-A-Million (%)
Direct logins	27.8	16.0	1.9	3.0	10.0
Google	40.2	31.7	17.9	20.8	45.9
Yahoo	5.9	4.9	6.3	3.5	1.9
Bing	15.1	5.8	1.3	2.9	0.0
Other search	2.9	3.3	0.0	0.0	1.9
Other bookstores	3.9	38.3	72.6	69.9	40.4
Price comparison	4.1	0.1	0.0	0.0	0.0

Source: comScore Search Planner data from August 2012.

related search terms. If this is the case, figure 5.2 may understate the relative importance of book-related searches on search engines. The importance of the long tail for book sales and searches remains an open question.

### 5.3.2 Incoming and Outgoing Traffic

Search engines are not the only way consumers reach a bookseller's site. Consumers may directly visit the site, be referred from another site (such as a price comparison site), and so on. Likewise, consumers that visit a bookseller's site may simply be gathering information before visiting another site to compare prices or make a purchase. This section examines these patterns of incoming and outgoing traffic.

#### *How Searchers Reach Booksellers*

To better understand how searchers actually reach the websites of different online booksellers, we analyze incoming traffic to five of the leading online bookstores in the United States (Amazon, Barnes & Noble, Books-A-Million, eCampus, and Powell's). In addition to observing the most popular search terms for these retailers, comScore's Search Planner data allows us to observe which sites get clicks for specific searches conducted at the main search engines as well as the source of incoming and the destination of outgoing traffic for each site.

The results summarized in table 5.4 show the percentages of visitors to these online bookstores in August 2012, categorized by the most important referring sites or platforms. For example, 27.8 percent of Amazon's traffic is due to direct logins, that is, consumers directly navigating to its site. Amazon received 40.2 percent of its traffic through Google, 5.9 percent through Yahoo, and 15.1 percent through Bing. As shown in figure 5.2, panel (a), most of this traffic stemmed from navigational searches, which are effectively a shortcut for direct logins. Finally, 3.9 percent of Amazon's traffic originated at another bookstore while 4.1 percent came from a price comparison site.

**Table 5.5** Percentage of outgoing visits from bookstores to search engines and other platforms

	Amazon (%)	Barnes & Noble (%)	eCampus (%)	Powell's (%)	Books-A-Million (%)
Logoff	43.0	31.4	5.7	29.8	27.9
Google	28.4	21.5	11.1	44.8	21.1
Yahoo	3.7	2.8	1.6	2.0	0.0
Bing	13.7	3.4	2.1	8.5	7.5
Other search	4.2	1.9	0.0	0.0	0.0
Other bookstores	4.2	38.7	79.5	14.8	43.5
Price comparison	2.9	0.3	0.0	0.0	0.0

Source: comScore Search Planner data from August 2012.

More broadly, table 5.4 shows that there is considerable heterogeneity in the sources of traffic for different booksellers. Amazon and Barnes & Noble receive much more traffic through direct logins than eCampus, Powell's, and Books-A-Million. This may stem from the fact that Amazon and Barnes & Noble are better-known brands; see Baye, De los Santos, and Wildenbeest (forthcoming). Powell's and Books-A-Million rely more heavily on general search engines to receive traffic, while referrals from other bookstores are relatively important for Barnes & Noble, eCampus, and Books-A-Million.<sup>16</sup> Finally, Amazon receives a far greater share of its traffic from price comparison sites than do the other booksellers.

#### *What Searchers Do After Visiting a Bookseller*

To identify what happens after having visited an online bookseller, we analyze the outgoing traffic from the bookstores using comScore's Search Planner data. The results are summarized in table 5.5, using data from August 2012 for Amazon, Barnes & Noble, eCampus, Powell's, and Books-A-Million.

First, note that a searcher that terminates his or her search activity after visiting a bookseller is more likely to have consummated a purchase than is a searcher that does not terminate but instead visits another bookseller or a price comparison site. Table 5.5 shows that 43 percent of users that visited Amazon terminated their search by logging off; this share is much lower for the other bookstores, especially for eCampus, which was 5.7 percent.

Second, note that consumers that choose to leave one bookstore and visit another bookstore may be doing so because the book was not in stock or was

16. The other bookstores category includes abebooks.com, alibris.com, amazon.com, audible.com, barnesandnoble.com, biggerbooks.com, bkstr.com, bncollege.com, bookbyte.com, booksamillion.com, campusbookrentals.com, christianbook.com, ebooks.com, ecampus.com, half.com, kobobooks.com, powells.com, and textbooks.com.

priced above their reservation price. Table 5.5 shows that only 4.2 percent of visitors to Amazon chose to visit another bookstore after searching there. This is significantly lower than for the other bookstores in our sample; for example, 79.5 percent of eCampus visitors chose to search another bookstore immediately after visiting the site.

Finally, consumers may visit a bookseller with the intent of free riding on the general information, reviews, or recommendations it provides, and then leave the bookseller's site to search for the best available price. As shown in table 5.5, 2.9 percent of Amazon visitors directly visit a price comparison site after viewing pages at Amazon. Visits to price comparison sites comprise a negligible share of outgoing traffic for the other booksellers in our sample.

### 5.3.3 Search Activity on Book-Oriented Platforms

As discussed earlier, general search engines are not the only way consumers can search for books. Consumers can search the websites of booksellers and other retailers, as well price comparison sites. In light of the fact that major booksellers receive a considerable share of their traffic from direct logins and navigational searches, one would expect the number of searches on these sites to be substantial. Unfortunately, comScore does not provide information about the keywords used for searches "inside" these retailers. Nonetheless, limited information is available through the qSearch database.

The qSearch data are based on the comScore panel, which contains all of the browsing activity of about two million users in the United States. It contains monthly data on searches conducted by consumers on the Internet.<sup>17</sup> These searches include searches through traditional search boxes as well as toolbars and widgets. The qSearch database includes the actual search volume and search intensity at over 200 online properties, including Amazon, Barnes & Noble, and several other book-oriented platforms. Unlike Search Planner, the data indicate the total number of searches on each site rather than a breakdown of searches by keyword. Thus, one cannot distinguish between searches for books versus other products with the qSearch data.

Table 5.6 provides a snapshot (February 2012) of the total number of searches conducted on the five largest book-oriented sites that are tracked by the qSearch database. Amazon processed the majority of searches among the bookstores, having almost ten times as many searches as the number two bookstore, Barnes & Noble. In terms of total searches, a smaller number of searches were conducted on the price comparison site (BookFinder). The

17. According to comScore's qSearch documentation, a search is defined as: (1) a user interaction where the user is presented with a search result page that contains results that match the consumer's search intent; (2) the search result page allows the user to refine or change the search parameters; and (3) the search can be initiated from a drop down menu or by clicking a link, as long as the first two rules are satisfied.

**Table 5.6** Number of searches on book-oriented platforms

	Searches (×1,000)	Searches (%)	Search visits (×1,000)	Search visits (%)	Searches per search visit
<i>Bookstores</i>					
Amazon	326,658	26.22	150,643	26.22	2.17
Barnes & Noble	37,205	2.99	10,620	2.99	3.50
AbeBooks	1,659	0.13	561	0.13	2.96
<i>Other book sites</i>					
Google Book Search	10,124	0.81	4,719	0.81	2.15
<i>Price comparison sites</i>					
BookFinder	692	0.06	220	0.06	3.15

Source: comScore qSearch, February 2012.

level of search activity is even more skewed toward Amazon when one measures searches using search visits.<sup>18</sup>

Figure 5.3 shows the evolution of searches between January 2011 and February 2013 for these book-oriented qSearch properties. The graph shows that, especially for Amazon and Barnes & Noble, there is seasonality in searches during the holiday shopping season, with searches increasing sharply near the end of the calendar year. Searches at Barnes & Noble appear to have peaked near the end of 2011, while searches at Amazon have generally increased, with the exception of the most recent few months.

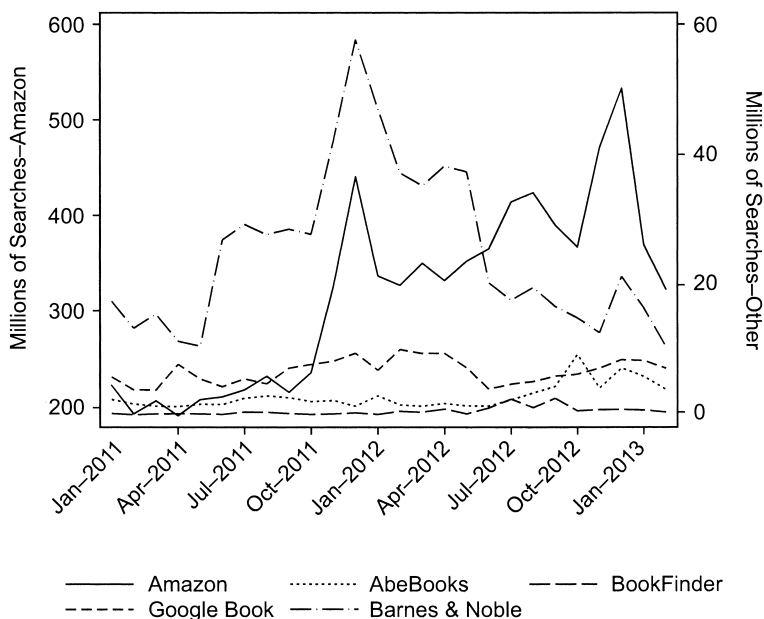
We emphasize that the qSearch data displayed in figure 5.3 does not indicate how many of the searches at these book-oriented platforms are, in fact, book related. While this is not so much an issue for sites that specialize in books (such as Barnes & Noble, BookFinder, and Google Book Search), it is clearly an issue for Amazon. While Amazon began its life as an online bookstore, it has evolved into a mass merchant where over 53 percent of its 2009 revenues derived from products outside the book category. And today, Amazon is itself a platform that services thousands of Amazon Marketplace sellers. Similarly, these data do not permit us to measure book-related searches at retailers such as Walmart or on platforms such as eBay (which received more than twice as many searches as Amazon in February 2012).

#### 5.4 The Dynamics of Book-Related Searches

This section provides a look at how searches for books on search engines evolve during a window around the release date of the book and also explores

18. According to comScore's qSearch documentation, a search visit is a session in which a user conducted one or more searches. If searches are conducted at different points during the day, with more than thirty minutes of search inactivity at the site, they count as multiple search visits.





**Fig. 5.3 Total searches at book-oriented sites**

*Source:* comScore qSearch, January 2011–February 2013.

*Note:* Amazon scaled on left, all others on right.

how the shifts toward other platforms have impacted book-related searches at the main search engines.

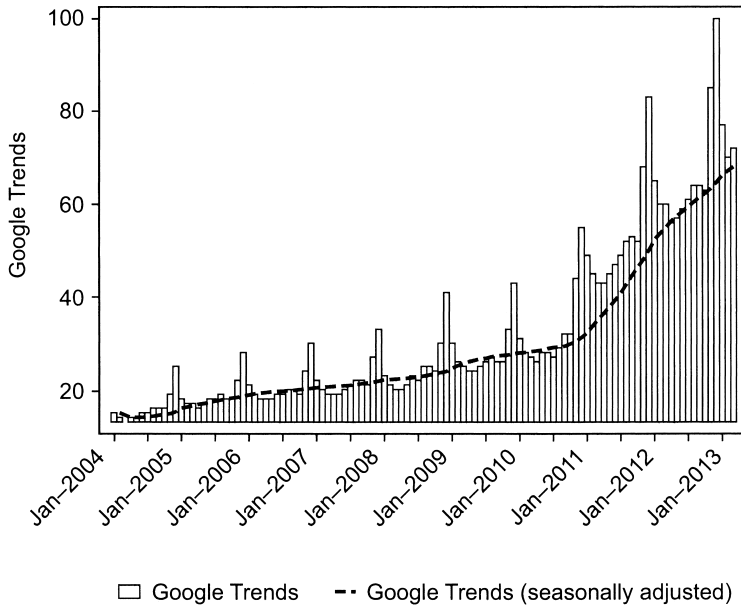
#### 5.4.1 Combining Data from comScore and Google Trends

Since our comScore database only contains information about searches for specific book-related keywords on a monthly basis for the 2010–2012 period, our strategy is to combine these data with data from Google Trends to create a data set that runs from 2004 through 2013.<sup>19</sup>

For a given search term, Google Trends provides a measure of the number of searches for that term between 2004 and 2013. The measure is relative: the maximum number of searches within the period is normalized to 100. To illustrate, figure 5.4 displays search volume reported by Google Trends for the term “Amazon.”<sup>20</sup> The bars represent monthly search “volume,” whereas the curve gives the seasonally adjusted trend. The figure shows search volume has been rising for the term “Amazon,” with the growth in

19. Choi and Varian (2012) describe how Google Trends data can help to predict economic time series of interest. Google Trends data has been used in epidemiology to detect influenza epidemics (Ginsberg et al. 2009) and in economics to predict the unemployment rate (D’Amuri and Marcucci 2010) and inflation (Guzmán 2011).

20. We limit search volume to the United States only.



**Fig. 5.4** Google Trends for the term “Amazon”

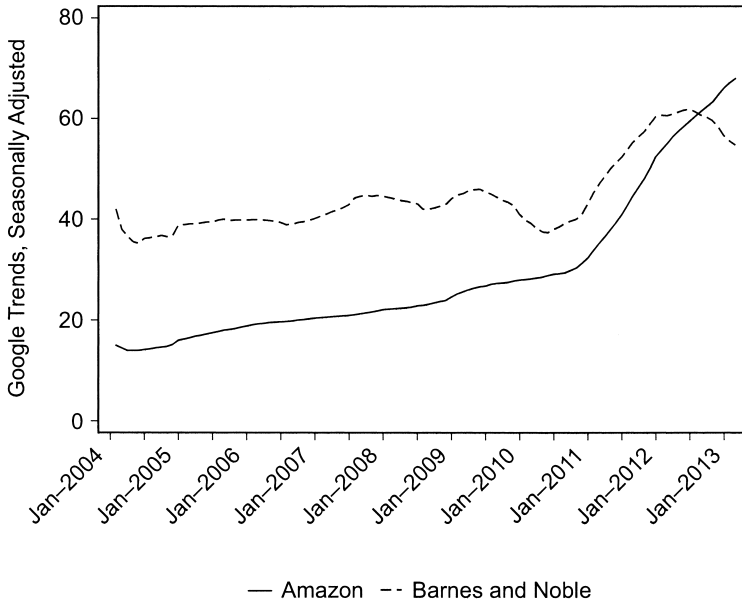
Source: Google Trends, January 2004–March 2013.

search volume accelerating after July 2010. Figure 5.5 displays the relative search volume over time for “Amazon” and “Barnes and Noble” in a single graph. The number of search queries for Barnes & Noble was relatively stable up to 2010, after which searches went up quickly. However, searches at Barnes & Noble peaked near the end of 2011, whereas the search volume for Amazon kept growing.

Since the Google Trends measure of search is relative, the relative position of the two lines in figure 5.5 is uninformative: it would be erroneous to infer that Barnes & Noble had more navigational searches than Amazon for the 2004–2012 period. Our strategy for overcoming this limitation is to use information about the actual number of searches for a given search term in comScore to convert the Google Trends index into historical data on the numbers of searches for that search term.<sup>21</sup>

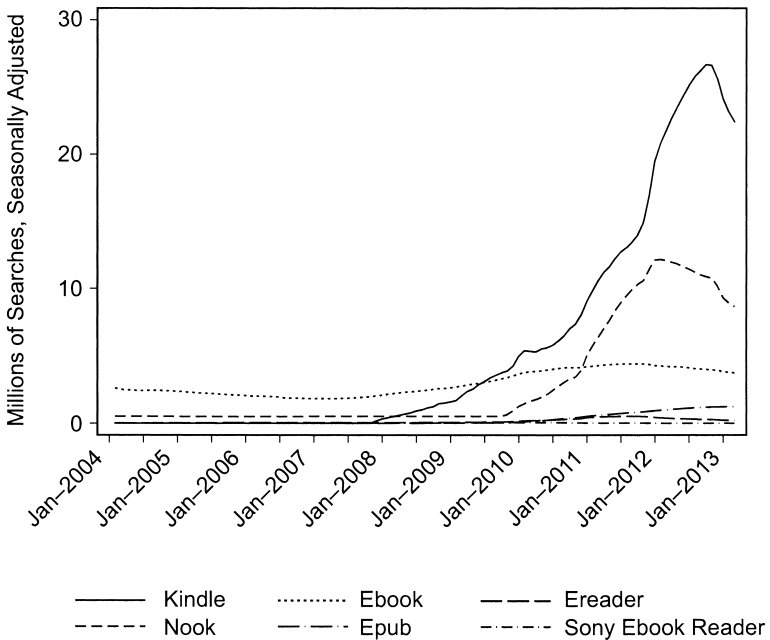
Figure 5.6 provides one example of the utility of combining Google Trends and comScore data. This graph shows the evolution of searches

21. Google Trends allows a comparison of up to five search terms. Nevertheless, the relative search volume of more than five search terms can be obtained by running multiple Google Trends comparisons, with each comparison having at least one search term that is common across all comparisons. Nonetheless, the comScore (or similar) data is still needed to pin down the level of searches. Also notice that searches according to the comScore data by definition lead to a click to a bookstore, whereas a search according to Google Trends may not result in a click.



**Fig. 5.5 Google Trends for the terms “Amazon” and “Barnes and Noble”**

Source: Google Trends, January 2004–March 2013.



**Fig. 5.6 E-reader-related searches on Google**

Source: Google Trends and comScore Search Planner.

for six terms related to e-readers: “kindle,” “nook,” “e-book,” “ereader,” and “sony ebook reader.” By combining the data, we not only obtain a time series dating back to 2004, but the resulting numbers of searches are cardinal so the different graphs may be compared. Figure 5.6 illustrates that searches for “kindle” took off in January 2008, and the number of searches for the Kindle has exceeded that for the Nook every year since. Notice that the recent decline in searches for these two e-readers is consistent with our earlier remarks regarding the projected decline in post-2012 e-reader sales as consumers switch to tablets and other devices.

#### 5.4.2 Searches for Specific Titles

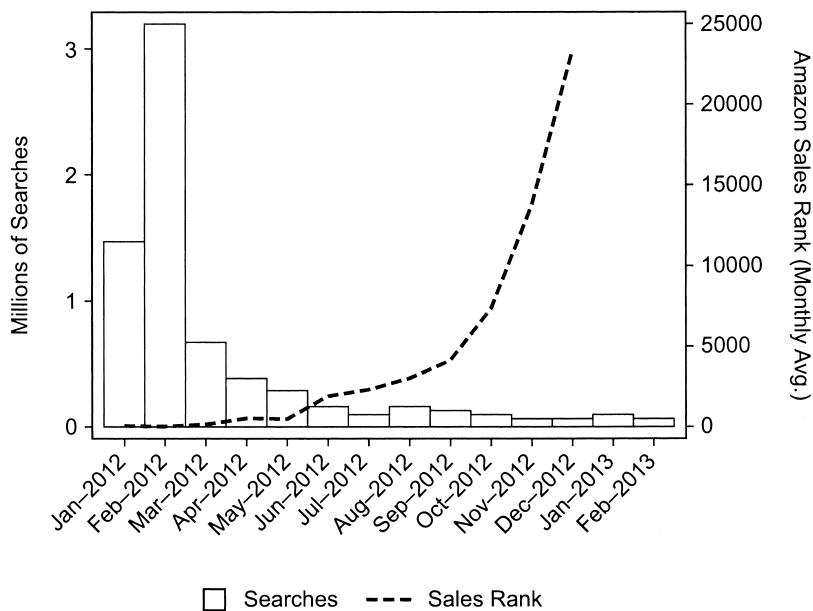
We are now in a position to take a deeper dive into searches for books on search engines and to examine how search patterns are evolving over time. We use the Search Planner data for a sample of 735 books to study how the number of clicks on Google changes throughout a book’s life cycle. Our sample includes titles released in every year since 2003, but roughly half of the sample is composed of books released in 2012. The search queries we use to obtain the clicks on Google are the exact titles of the book in our sample. In addition to clicks, we observe book characteristics such as ISBN, list price, publisher, date of publication, format (hardcover or paperback), number of pages, edition (e.g., first or reprint), dimension, and weight. Moreover, for 2012 we have daily observations of a book’s price as well as availability at each of the leading online bookstores.<sup>22</sup> For Amazon we also observe a book’s customer rating, number of customer reviews, and sales rank. Most books in our sample are first editions; the vast majority are *New York Times* bestsellers that came out during 2011 and 2012.

The Search Planner data has two limitations. The first limitation is that we only observe clicks for the period between February 2010 and February 2013. The second is that we only observe clicks if they exceed a certain threshold. To deal with these two matters, we supplement the Search Planner data with Google Trends data, as discussed above. For each book in our sample we use Google Trends to capture the volume of search queries that users enter into Google when searching for the title of a book.<sup>23</sup> Since Google Trends only gives a relative measure of search volume, we convert the index to actual levels of search volume by setting the maximum value of the index during the period between February 2010 and February 2013 equal to the corresponding clicks according to Search Planner.

Figure 5.7 gives an example for the book *The End of Illness* by David B. Agus, which was released on January 17, 2012. The number of searches on Google for this title is relatively high shortly after release, but within a few months declines to a relatively low level. Also shown in this graph is Ama-

22. These are amazon.com, barnesandnoble.com, booksamillion.com, walmart.com, powells.com, and ecampus.com. See section 5.5 for a more detailed description of the price data.

23. We use exact phrases (exact words in quotes), for example, “Gone girl” instead of Gone girl. We limit search volume to the United States only.



**Fig. 5.7** Google searches for *The End of Illness*

Source: comScore Search Planner.

zon's sales rank for this title, which shows a reverse pattern. This illustrates that sales rank is actually inversely related to the number of searches. It is an open question whether this is a result of lags in Amazon's construction of sales rank or whether this means that, during the life cycle of a book, consumers increasingly visit Amazon's site directly to purchase the book rather than using a general search engine.

The pattern for this particular title is fairly representative of the other titles in our sample. In particular, we ran a linear regression of the total number of searches for each book title in our sample on a constant and a set of dummies that represent months since release. The underlying regression results for various specifications are shown in table 5.7. For the specification in column (1), which includes book fixed effects to control for heterogeneity among books, most of the month-since-release coefficients are significantly different from zero at conventional levels. Figure 5.8 plots the predicted values of each of the dummies using the parameter estimates for specification (1). This figure shows that the patterns of search observed around the release window for *The End of Illness* (see figure 5.7) is typical of newly released books: search activity tends to be high in the first few months, but then quickly drops to significantly lower levels.

Column (2) of table 5.7 is a specification that includes book characteristics rather than book fixed effects. Although this reduces the fit of the model and

**Table 5.7** Regression results

	(1)		(2)		(3)	
<i>Months before release</i>						
2 months	-0.696	(0.188)***	-0.688	(0.281)**	-0.765	(0.282)***
1 month	-0.734	(0.188)***	-0.732	(0.281)***	-0.771	(0.281)***
<i>Months after release</i>						
1 month	-0.299	(0.188)	-0.297	(0.281)	-0.257	(0.281)
2 months	-0.426	(0.188)**	-0.424	(0.281)	-0.342	(0.282)
3 months	-0.540	(0.188)***	-0.537	(0.281)*	-0.413	(0.282)
4 months	-0.503	(0.188)***	-0.499	(0.281)*	-0.329	(0.284)
5 months	-0.540	(0.191)***	-0.515	(0.285)*	-0.311	(0.288)
6 months	-0.654	(0.195)***	-0.628	(0.291)**	-0.391	(0.296)
7 months	-0.681	(0.199)***	-0.581	(0.297)*	-0.318	(0.303)
8 months	-0.692	(0.202)***	-0.580	(0.301)*	-0.280	(0.308)
9 months	-0.725	(0.205)***	-0.612	(0.306)**	-0.277	(0.315)
10 months	-0.747	(0.209)***	-0.613	(0.312)**	-0.248	(0.323)
11 months	-0.798	(0.216)***	-0.600	(0.321)*	-0.212	(0.333)
12 months	-0.719	(0.221)***	-0.510	(0.328)	-0.095	(0.342)
First edition indicator			-0.813	(0.200)***	-0.820	(0.202)***
<i>Format</i>						
Mass market paperback			1.228	(0.249)***	1.467	(0.255)***
Paperback			0.757	(0.208)***	0.736	(0.208)***
Other			0.256	(0.365)	0.660	(0.377)*
Pages			-0.005	(0.000)***	-0.005	(0.001)***
Weight			0.097	(0.007)***	0.099	(0.007)***
Time index					0.094	(0.024)***
Time index squared					-0.001	(0.000)***
Constant	1.398	(0.133)***	1.049	(0.297)***	-1.344	(0.913)
Book indicators		Yes		No		No
Publisher indicators		No		No		Yes
Observations		10,615		10,585		10,585
R-squared		0.664		0.213		0.214

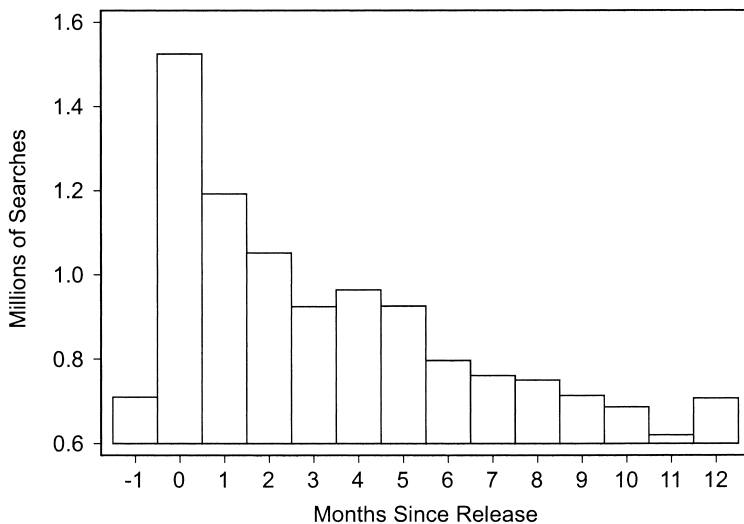
*Notes:* Dependent variable: number of searches for a specific book title. Standard errors in parentheses.

\*\*\*Significant at the 1 percent level.

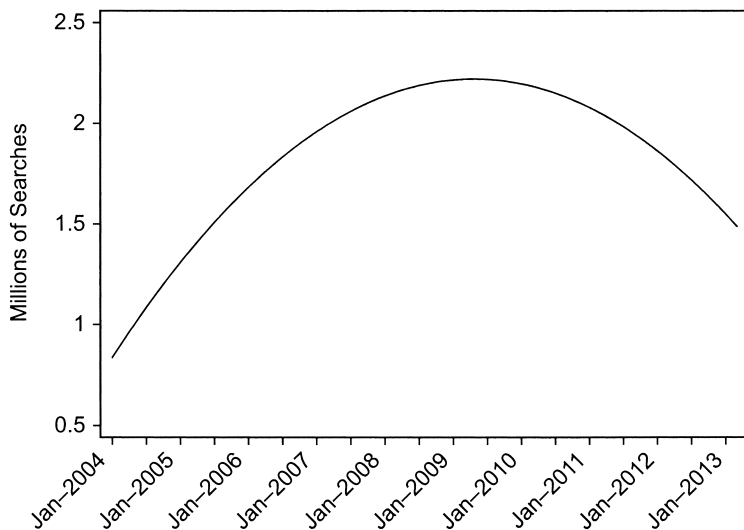
\*\*Significant at the 5 percent level.

\*Significant at the 10 percent level.

the significance of some of the months-since-release dummies, the magnitude of the coefficients is similar to that in the first specification. Most book characteristic variables have parameter estimates that are significantly different from zero. There is less search for first edition books than for books in later editions. This may stem from the fact that later editions are only published if the book is popular. There are more searches for paperbacks than hardcover books in our sample, and again this may reflect the fact that paperbacks are targeted to the masses and therefore result in more searches. The final column of table 5.7 adds a linear and quadratic time trend to control for possible changes in search behavior between 2004 and 2013. Both



**Fig. 5.8** Predicted searches for book titles on Google



**Fig. 5.9** Predicted searches per book title on Google, 2004–2013

coefficients are highly significant, suggesting that (controlling for changes in search behavior around release windows, edition, format, and so on) the number of book searches at search engines is evolving.

To more clearly see the implied evolution of searches for books over time, figure 5.9 plots the number of searches predicted by the model for each year in our sample. This plot shows that searches per book on Google increased

between 2003 and 2007, but since then have declined. This predicted pattern is consistent with our view that searchers are shifting away from using general search engines to find books and are increasingly (1) visiting online bookstores (such as Amazon and Barnes & Noble) to conduct searches, and (2) searching within closed systems (e.g., Kindle or Nook) to purchase books. Thus, while we cannot directly measure book-related searches within these alternative search platforms, the observable data from search engines is consistent with this hypothesis.

We conclude by noting that the pattern in figure 5.9 is also consistent with the trend in average monthly retail sales for brick-and-mortar bookstores, as shown in figure 5.1, and may reflect the same structural changes in the book industry that have led to a decline in the sales of physical books at traditional bookstores. For instance, the decline coincides with the increasing popularity of e-readers shown earlier in figure 5.6, the general shift toward e-books, and the rising prominence of Amazon as a platform for searching for books and other products.

## 5.5 Book Prices

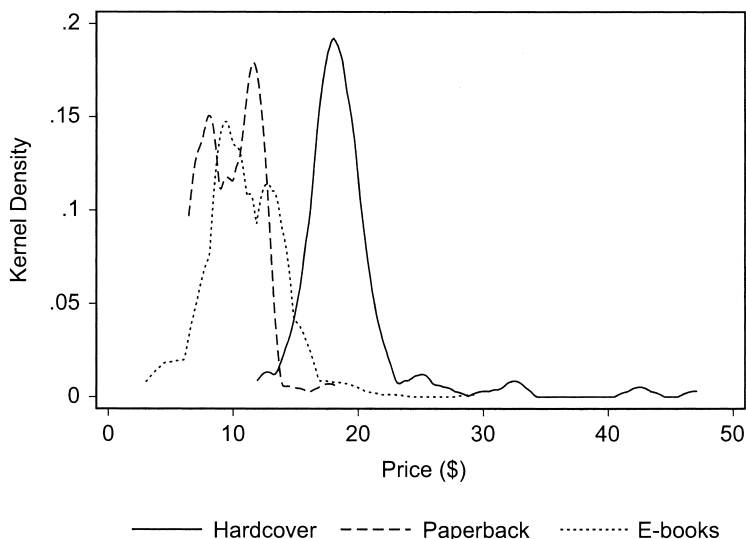
Another open question is how changes in search technologies and behavior have impacted the prices of physical and digital books. A priori, it is not clear how these prices should evolve over time. In the short run, intermediaries may use penetration pricing in an attempt to induce consumers to switch from buying printed books to e-books. Once the market is more mature, and a sufficient number of consumers is locked into a specific platform, firms may find it optimal to increase prices. Sales of physical and digital books are also likely to be related. Hu and Smith (2013) use a natural experiment to analyze the effect of e-book availability on sales of physical books and find that delayed e-book availability results in a large decrease in total e-book sales.

We analyze book prices using a large data set of daily prices collected using a scraper written in Java. The data set contains daily price information throughout 2012 for physical books as well as e-books sold at Amazon, Barnes & Noble, and Books-A-Million, physical books sold at eCampus, Powell's Books, and Walmart, and e-books sold at Kobo and Google. In addition to prices, we observe availability and several book characteristics such as the publisher, format (hardcover or paperback), edition, number of pages, weight, and dimension.

Figure 5.10 shows the kernel estimates of the price densities for hardcover books, paperbacks, and e-books. The prices used are average prices across books within format and across bookstores, so the variation reflects changes over time. E-books are on average less expensive than hardcover books, but slightly more expensive than paperbacks.

The relationship between publishers and retailers complicates the anal-





**Fig. 5.10 Price distributions for hardcover, paperback, and digital books**

*Note:* Data truncated at \$50.

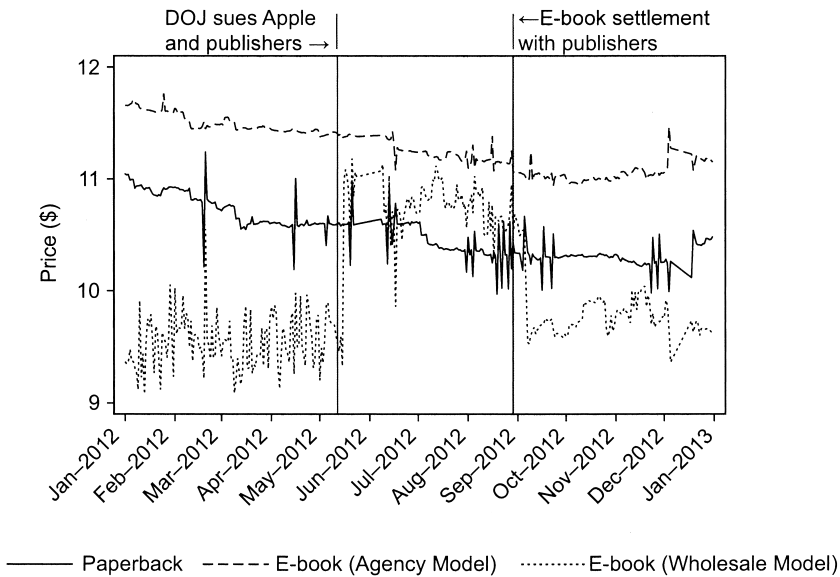
ysis of prices for e-books. Traditionally, books were sold using a wholesale model: a publisher sells a book to a retailer at a wholesale price, which is typically half the list price for the book. The retailer is then free to sell the book at whatever price it wants. Initially this model was used by all players for selling e-books as well. However, to foster e-book sales, Amazon began heavily discounting popular e-books, selling them for \$9.99—a price that is below the wholesale price of many physical books. Publishers, fearing that e-books would cannibalize sales of printed books and that consumers would grow accustomed to low e-book prices, were concerned about the low e-book prices. When Apple launched its iBookstore in April 2010, five of the “big six” publishers (HarperCollins, Hachette, Macmillan, Penguin, and Simon & Schuster) adopted Apple’s newly developed agency model. Under the agency model the retailer receives a fixed percentage of the book’s sale price (typically 30 percent). Moreover, under the agency model, the retail price is no longer set by the retailer but set by the publisher. Other online bookstores reached similar agreements with the publishers. As a reaction to these developments, the Department of Justice sued Apple and the five publishers on April 12, 2012, alleging that they conspired to fix the prices of e-books.

Table 5.8 provides descriptive statistics for the prices of books in our sample. The table makes a distinction between e-books that are sold under the agency model and those that are sold under the wholesale model. Average prices are generally higher for the agency model e-books and show less variation.

Figure 5.11 plots the average prices of paperbacks, e-books sold under

**Table 5.8** Descriptive statistics for book prices

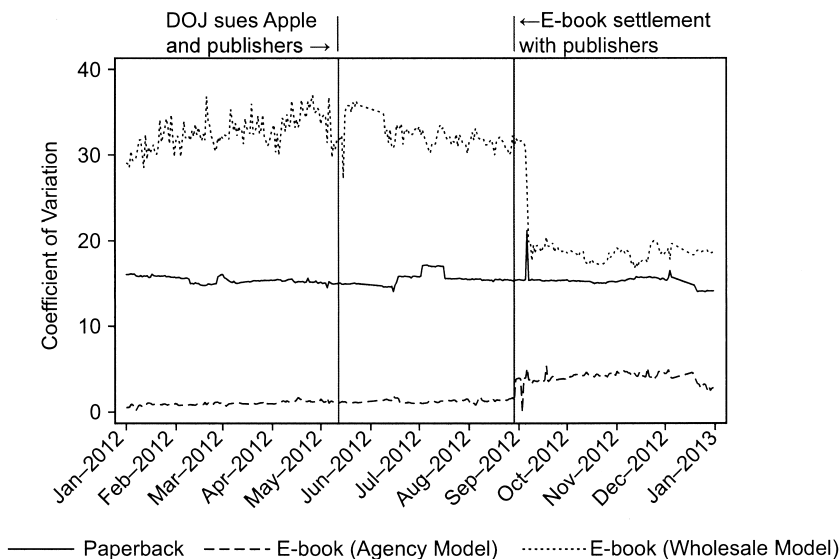
Bookstore	Hardcovers		Paperbacks		E-books (agency)		E-books (wholesale)	
	Mean (\$)	Std. dev.	Mean (\$)	Std. dev.	Mean (\$)	Std. dev.	Mean (\$)	Std. dev.
<i>Both formats</i>								
Amazon	16.3	5.7	9.9	6.3	11.2	2.7	8.6	7.5
Barnes & Noble	16.8	6.0	10.5	7.1	11.3	2.6	10.2	5.0
Books-A-Million	16.7	6.6	10.2	9.7	11.4	2.5	9.9	4.2
<i>Physical books only</i>								
eCampus	21.9	4.9	10.5	9.0				
Powell's	27.0	7.1	12.9	4.2				
Walmart	16.2	3.6	9.0	7.3				
<i>E-books only</i>								
Kobo					11.1	2.6	11.0	6.6
Google					11.1	3.4	9.8	8.1



**Fig. 5.11** Average book prices by book format

the agency model, and e-books sold under the wholesale model across bookstores.<sup>24</sup> Even though the DOJ reached a settlement with three of the publishers in April 2012 (HarperCollins, Hachette, and Macmillan), and as part of the settlement e-book retailers are allowed to discount e-book titles, prices for e-books sold under the wholesale model increased shortly after the

24. Since we do not have sales data, average prices are not weighted by sales.



**Fig. 5.12** Average coefficient of variation by book format

lawsuit was announced. The steep decline in e-book prices sold under the wholesale model a few months later coincides with the final approval of the settlement between the DOJ and the three publishers, with HarperCollins books starting to be discounted on Amazon soon after (Hachette and Simon & Schuster discounting began in December 2012).

One might wonder whether the evolution of search has impacted the degree of price dispersion for books. One standard measure of dispersion is the coefficient of variation—the ratio of the standard deviation in the prices different sellers charge for a given book divided by the average price sellers charge for that book. Since this measure is homogeneous of degree zero in the level of prices, one can average these coefficients of variation across all books in our sample to obtain a useful summary statistic for the level of price dispersion. Figure 5.12 plots the average coefficient of variation over time for the three book formats in our sample. Notice that the coefficient of variation is highest for e-books sold under the wholesale model, lower for paperback books, and lowest for e-books sold under the agency model. Additionally, note that the coefficient of variation for e-books sold under the wholesale model declined significantly after the e-book settlement, while the coefficient of variation for e-books sold under the agency model increased slightly. In contrast, the coefficient of variation for paperbacks was relatively stable throughout the entire period.

The underlying reason for these changes in prices and levels of price dispersion remains an open question. We note that the two remaining publishers (Macmillan and Penguin) settled in early 2013, so this event is unlikely to be the driver of the patterns observed in figures 5.11 and 5.12.

## 5.6 Concluding Remarks

Our initial look at the evolution of platforms through which consumers search for, purchase, and read books highlights a number of issues that should be on the agenda of those interested in digital media and other products sold online. There are a variety of different platforms where consumers can search for books, including online bookstores (such as Barnes & Noble), mass merchants (Amazon), and price comparison sites (BookFinder). While data suggest that consumers are increasingly conducting searches on closed systems (such as e-readers) and retailer platforms (such as Amazon), measurement of these trends is difficult. In contrast to these platforms, data on the types of searches conducted on search engines is readily available.

While Amazon and Barnes & Noble get significant numbers of visits from those conducting searches on Google and Bing, most of the observed searches are navigational ones—situations in which users query a search engine as a shortcut for navigating to the site of the bookseller where actual product searches are performed. Navigational searches to these top booksellers have steadily increased since 2004. The number of browser-based searches on book-related platforms such as Amazon, Barnes & Noble, BookFinder, and Google Books is sizable, but it is an open question whether browser-based searches on these platforms will continue to grow. Indeed, our preliminary analysis of searches for book titles on Google suggests a declining trend in book searches since 2007. While these patterns are consistent with shifts toward nonbrowser searches on closed platforms, including smartphone and tablet apps, additional research is required to reach any definitive conclusions.

Another open agenda item is the importance of the long tail for product searches in general, and more specifically for book searches. The search terms corresponding to long-tail searches at the major search engines are not observed in comScore data because of low search volume; we do not observe niche searches, which presumably take place in the long tail. Additional research will be needed to analyze the long tail.

An interesting new area of research has emerged relating to e-books. We have identified several new developments with respect to pricing strategies of publishers and booksellers and have provided some preliminary findings on how some of these developments have influenced book prices. More research will be needed on how different selling formats (e.g., the agency and wholesale models) affect pricing more generally, from both a theoretical and empirical point of view.<sup>25</sup>

Finally, note that we have completely sidestepped the first-order question of what consumers are seeking when they conduct searches for books, digital media, or (more broadly) other products. More bluntly, this chapter

25. Johnson (2013) and Abhishek, Jerath, and Zhang (2013) develop theoretical models that are geared toward a comparison of the different selling formats in the book industry.

says nothing about whether consumers are searching for information about the best book to read or information about the best prices available. Future research that unpacks these and other motives for search would be a valuable addition to the literature.

## References

- Abhishek, Vibhanshu, Kinshuk Jerath, and Z. John Zhang. 2013. "Agency Selling or Reselling? Channel Structures in Electronic Retailing." Working Paper, Carnegie Mellon University.
- Baye, Michael R., Babur De los Santos, and Matthijs R. Wildenbeest. 2012. "What's in a Name? Measuring Prominence, and Its Impact on Organic Traffic from Search Engines." Working Paper no. 2012-09, Kelley School of Business, Indiana University. <http://dx.doi.org/10.2139/ssrn.2191051>.
- . 2013. "The Evolution of Product Search." *Journal of Law, Economics & Policy* 9 (2): 201–21.
- . Forthcoming. "Search Engine Optimization: What Drives Organic Traffic to Retail Sites?" *Journal of Economics and Management Strategy*.
- Baye, Michael R., John Morgan, and Patrick Scholten. 2006. "Information, Search, and Price Dispersion." In *Handbook in Economics and Information Systems*, edited by T. Hendershott. Amsterdam: Elsevier.
- Brynjolfsson, Erik, Yu (Jeffrey) Hu, and Michael D. Smith. 2003. "Consumer Surplus in the Digital Economy: Estimating the Value of Increased Product Variety at Online Booksellers." *Management Science* 49 (11): 1580–96.
- . 2011. "Goodbye Pareto Principle, Hello Long Tail: The Effect of Search Costs on the Concentration of Product Sales." *Management Science* 57 (8): 1373–86.
- Chevalier, Judith, and Austan Goolsbee. 2003. "Measuring Prices and Price Competition Online: Amazon.com and BarnesandNoble.com." *Quantitative Marketing and Economics* 1 (2): 203–22.
- Chevalier, Judith A., and Dina Mayzlin. 2006. "The Effect of Word of Mouth on Sales: Online Book Reviews." *Journal of Marketing Research* 43 (3): 345–54.
- Choi, Hyunyoung, and Hal Varian. 2012. "Predicting the Present with Google Trends." *Economic Record* 88 (s1): 2–9.
- Clay, Karen, Ramayya Krishnan, and Eric Wolff. 2001. "Prices and Price Dispersion on the Web: Evidence from the Online Book Industry." *Journal of Industrial Economics* 49 (4): 521–39.
- D'Amuri, Francesco, and Juri Marcucci. 2010. "Google It! Forecasting the US Unemployment Rate with Google Job Search Index." Working Paper no. 18732, Munich Personal RePEc Archive.
- De los Santos, Babur. 2012. "Consumer Search on the Internet." Working Paper no. 08-15, NET Institute. <http://dx.doi.org/10.2139/ssrn.1285773>.
- De los Santos, Babur, Ali Hortaçsu, and Matthijs R. Wildenbeest. 2012. "Testing Model of Consumer Search Using Data on Web Browsing and Purchasing Behavior." *American Economic Review* 102 (6): 2955–80.
- De los Santos, Babur, and Matthijs R. Wildenbeest. 2014. "E-book Pricing and Vertical Restraints." Working Paper, Indiana University.

- Forman, Chris, Anindya Ghose, and Avi Goldfarb. 2009. "Competition between Local and Electronic Markets: How the Benefit of Buying Online Depends on Where You Live." *Management Science* 55 (1): 47–57.
- Ginsberg, Jeremy, Matthew H. Mohebbi, Rajan S. Patel, Lynnette Brammer, Mark S. Smolinski, and Larry Brilliant. 2009. "Detecting Influenza Epidemics Using Search Engine Query Data." *Nature* 457 (7232): 1012–14.
- Guzmán, Giselle. 2011. "Internet Search Behavior as an Economic Forecasting Tool: The Case of Inflation Expectations." *Journal of Economic & Social Measurement* 36 (3): 119–67.
- Hong, Han, and Matthew Shum. 2006. "Using Price Distributions to Estimate Search Costs." *RAND Journal of Economics* 37 (2): 257–75.
- Hu, Yu (Jeffrey), and Michael D. Smith. 2013. "The Impact of e-Book Distribution on Print Sales: Analysis of a Natural Experiment." Working Paper, Carnegie Mellon University.
- Johnson, Justin. 2013. "The Agency and Wholesale Models in Electronic Content Markets." Working Paper, Cornell University.
- Manley, Laura, and Robert P. Holley. 2012. "History of the e-Book: The Changing Face of Books." *Technical Services Quarterly* 29 (4): 292–311.
- Moraga-González, José Luis, and Matthijs R. Wildenbeest. 2012. "Comparison Sites." In *Handbook of the Digital Economy*, edited by Martin Peitz and Joel Waldfogel, 224–53. Oxford: Oxford University Press.
- Rysman, Marc. 2009. "The Economics of Two-Sided Markets." *Journal of Economic Perspectives* 23 (3): 125–43.
- Stigler, George. 1961. "The Economics of Information." *Journal of Political Economy* 69 (3): 213–25.

## Comment      Marc Rysman

The book publishing industry has provided one of the most visible and tangible digital transformations of market. Most of the focus has been on the production, distribution, and ownership rights in the context of digital e-books, or on how online booksellers compare to offline booksellers. However, little attention has been paid to how consumers find the books that they ultimately buy. Browsing bookstores, asking store employees for suggestions, and waiting for reviews in your local newspaper has been replaced by online opinions, suggestion engines, and immediate access to an enormous number of outlets that review books. Even in the case where a consumer already knows the book they want and they merely have to locate the book, digitization has important implications. It is much easier to search across a broad set of sellers on the Internet—indeed, shopping websites specifically facilitates this activity. The chapter by Baye, De Los Santos, and Wildenbeest

Marc Rysman is professor of economics at Boston University.

For acknowledgments, sources of research support, and disclosure of the author's material financial relationships, if any, please see <http://www.nber.org/chapters/c13126.ack>.