

## The Everything Store?

Once Amazon started to build its extensive warehouse network, Bezos envisioned “Project Fargo,” a reference to the 1996 Coen brothers’ film, where the company would have at least one of every product in stock.<sup>1</sup> Early Lab126 mission statements echoed the sentiment for ebooks, with the desire to have a copy of every book ever published available for download.<sup>2</sup> The US Kindle Store boasted over 5.5 million titles in September 2017, but Amazon faced an early hurdle to offer a more modest catalog of 88,000 titles at launch. The technical barriers of digitization were insignificant compared to the social challenges involved in persuading publishers to create ebooks. Publishers were comfortable with producing books digitally, but they remained skeptical of digital distribution before the Kindle’s launch. The gamble paid off, and Amazon boasted a large catalog in November 2007 “including more than 101 of 112 current *New York Times* Best Sellers.”<sup>3</sup> The combination of best sellers and a range of older titles built a critical mass for the Kindle’s launch and encouraged consumers that publishers were serious about ebooks.

Ebook publishing divides into four main segments: *frontlist*, *backlist*, *public domain*, and *digital-only / self-published* works. The frontlist and backlist are mainstays of legacy publishing. Publishers release new titles, called the frontlist, aiming to sell them in high quantities and create a best seller. After the peak window for sales, thought to be as short as twelve weeks for most books, titles move to the backlist, where they remain in print if they continue to sell enough copies to justify a reprint. For publishers with a rich history, such as Penguin, Faber & Faber, and Oxford

University Press, the backlist represents a valuable revenue stream, as readers continue to consume historically significant titles still protected by copyright. As of 2020, books published before 1924 are in the public domain in the United States and can be published without prior permission or for free on the Kindle. Ebook platforms' infrastructure bypasses the workflows and timescales of legacy publishing. Start-ups have taken advantage of these new freedoms, including digital-only publishers such as Canelo and self-published authors who choose to circumvent the gate-keeping function of publishers entirely. Recent journalism prophesying the demise of ebooks focuses on print-oriented publishers' difficulties in adapting to the demands of digital publishing with new titles, and the expenses of digitizing the backlist, while ignoring the rise of digital-only publications owing to their association with vanity presses.<sup>4</sup> There are a few exceptions: Mills & Boon, a romance publisher, has pivoted toward digital publications to satiate its audience's demands. Other publishers focus on print first with ebooks remaining an afterthought, outside of a few publishers who continue to hire digital-publishing advocates. The long-term success of ebook platforms such as the Kindle depends on a fine balance of these four segments. Kindle Direct Publishing and the promotional value of Kindle Unlimited built new audiences for born-digital content, but this increased readership exists in a symbiotic state with more traditional publishing, which drives readers to the Kindle Store before discovering new genres and experiences.

Despite initial optimism, print-oriented publishers now contest any challenge to the orthodoxy. For example, in 2008 Frankfurter Buchmesse participants agreed that digital sales would overtake print by 2018, and Dick Brass, vice president of technology development for Microsoft, predicted that "the last paper edition of the *New York Times* will appear in 2018."<sup>5</sup> Nonetheless, popular Kindle titles still correlate to the whims of print. From June 2010 to November 2011, Amazon maintained the Kindle Million Club, a list of authors who had sold over a million ebooks.<sup>6</sup> In the eighteen months this list was active, fourteen members were inducted. Stieg Larsson was the first inductee in July 2010, followed by other fiction heavyweights such as James Patterson, Lee Child, Suzanne Collins, and Michael Connelly in the following twelve months. Two Million Club members, John Locke and Amanda Hocking, were self-published, offering an early indication of the changes afoot. Contrary to the perceived phenomenon of "mommy porn" in the years since the release of *Fifty Shades of Grey*, this list reveals a convergence toward crime fiction and thrillers, with the occasional science fiction and young adult title. Many brand-name authors only entered the club with the release of a film or television adaptation of

their works. For example, George R. R. Martin sold over a million copies of his works after the popularity of the first season of *Game of Thrones* in September 2011, and Stephenie Meyer's appearance on the list in November 2011 coincides with the penultimate *Twilight Saga's* release date in cinemas.

The club has several notable absences, including heavyweight authors such as Stephen King and J. K. Rowling. Both decided against publishing ebooks owing to skepticism about the format. King's hesitance was born out of the failed attempt to serialize *The Plant* in 2000, although his books had continued to be available for sale as ebooks, on Amazon no less, before the Kindle's launch. The success of the first-generation device persuaded King to pen a promotional story, "UR," released with a cover adorned with his custom pink Kindle 2, to mark the launch of the device. King joined Bezos for the device's launch in New York City's Morgan Library in February 2009, fully endorsing the device. Nonetheless, in the two years after King's entrance to the Kindle market, he had not sold a million copies despite having forty novels in print at the time. Rowling was another high-profile abstainer and resisted the Kindle's siren call until 2011, when she launched Pottermore to challenge Amazon's ebook dominance. Pottermore was initially designed as an interactive version of the seven core *Harry Potter* books featuring several gamelike features but evolved into the exclusive ebook distributor for the series. Rowling elected to distribute ebooks directly via Kindle Unlimited in 2015 to promote the subscription service. Readers who wished to purchase the titles were instructed to visit Pottermore, allowing Rowling to sidestep the Amazon infrastructure and dictate terms of access. Rowling's position was unique, as the *Harry Potter* series had greater cultural and economic clout than other authors, and she could dictate terms to Amazon: Nielsen BookScan data show that Rowling sold 1 percent of all books sold in the United Kingdom between 2001 and 2018. Rowling and King raised the potential for an alternative ebook ecosystem, disconnected from the whims of Amazon, but as the market has consolidated, even the refuseniks have fallen into line. Without a concerted effort from major publishers and high-profile authors, it is unlikely that Amazon's grip on the ebook market will slip.

### The Frontlist and Best Sellers

The best seller is an enigma in publishing studies, as it is one of the only metrics embraced by an industry that otherwise shuns quantification. Best sellers are editorialized and depend on algorithms with clear human biases. For example, when the *New York Times* publishes its best-seller

list, the data are collected from a range of self-reporting sources filtered by “standards for inclusion that encompass proprietary vetting and audit protocols, corroborative reporting and other statistical determinations.”<sup>7</sup> The *Times* keeps the exact list of bookshops secret, to discourage mass purchasing from suspected reporting stores, but its precautions do not stop authors from gaming the system. Lani Sarem’s young adult novel *Handbook for Mortals* was dropped from the best-seller list in August 2017 when bookstores reported suspicious bulk-buying patterns in an early draft of the list circulated to trade members.<sup>8</sup> Further levels of curation exist. Cookbooks often outsell other trade nonfiction but remain conspicuously absent from best-seller lists. The newspaper also created a separate list for children’s books after the *Harry Potter* series became a perennial fixture on the trade fiction list. Amazon challenged this orthodoxy from its early categorization of best sellers. Readers could find the most popular title within microgenres, with the expectation of real-time updates rather than offering a snapshot of the previous week.

Amazon adopted its familiar strategy of discounting best sellers to attract users to the Kindle. *New York Times* Best Sellers available at launch were capped at \$9.99, leading to discounts as high as \$20 for hardcover nonfiction titles to establish trust in a new format and counter early skepticism toward buying digital books. The value of a digital version of a newly published hardcover remains hotly contested by major publishers and Amazon over the last decade after this initial reduction in value despite publishers’ earlier trials of the scheme, including Random House’s launch of \$9.99 titles for the AtRandom imprint in 2000.<sup>9</sup> After these skirmishes, agency pricing is the current paradigm for big publishers. Agency pricing fixes an ebook’s cost, with a note from Amazon that the price was publisher mandated. Under this model, “the retailer acts as an agent of the publisher, which itself sets the retail price of the ebooks, with the retailer taking a commission” of 30 percent.<sup>10</sup> Titles protected by agency pricing can cost more than their print equivalents. For example, Philip Pullman’s *La Belle Sauvage* had a recommended retail price of £20 but was available in hardcover for £9 and Kindle edition for £9.99 when published in October 2017. Likewise, Elena Favilli and Francesca Cavallo’s *Good Night Stories for Rebel Girls* (£16.99 RRP) was £8.49 in print but £9.99 for the Kindle.<sup>11</sup> Ebooks in the United Kingdom were subject to a 20 percent value-added tax until May 2020, while print is exempt, accounting for the higher price, but the distribution costs of ebooks are substantially lower.<sup>12</sup> In recent years, the industry has therefore favored price matching to ensure a robust hardcover market by fixing the prices of ebooks. The higher cost of goods clashes with the public’s perception that an ebook should be relatively

inexpensive compared to the print edition. Increasing the costs of new books shrank the potential size of the market, but ebooks could still be useful for developing the backlist. Conversely, the gold rush of Kindle Direct Publishing (KDP) was predicated on cheap titles, hollowing out the potential for titles that sit between cheap digital-only publications and the more expensive releases of print-oriented publishers.

### Digitizing the Backlist

The backlist presented a greater opportunity for publishers that might have a treasure trove of older titles that might be profitable if sold at low cost in high volume. Although publishers had used digital tools to produce books for decades, material created before the introduction of digital workflows remained inaccessible. Since the 1980s, early innovators such as Coach House Press used Standardized General Markup Language (SGML), the precursor to HTML, to aid typesetting. This practice is now widespread, and print books start their lives as “born-digital” objects, easing the transition from print to ebook versions.<sup>13</sup> Titles published before the 1990s were difficult to convert to an appropriate format, as no digital typescript existed for many titles that had not enjoyed enough success to merit a reprint.

The public domain was a known quantity. Organizations such as Project Gutenberg, Early English Books Online, and the Oxford Text Archive started digitizing public domain books in the early 1970s. The projects’ investment in digitizing public domain titles created opportunities for publishers to produce ebooks for a low price or free, which led to several start-ups focusing on public domain titles, including Aegitas, Bartleby.com, and Enhanced Media Publishing.<sup>14</sup> Public domain ebooks are low cost with little risk, as any further editorial work is optional, and these works come with a preformed market. Jacob Flynn, Rebecca Giblin, and François Petitjean’s analysis of copyright terms reveals that books in the public domain receive more investment than those under direct ownership of a single publisher, especially as competing ebooks can be published, thus incentivizing higher quality.<sup>15</sup>

Amazon removed editorial oversight from the publication of public domain titles, and as a result, there has been an increase in poor-quality editions that have not been edited to clean up the scanned text or otherwise create a high-quality edition. For example, Whitney Trettien has traced a zombie print-on-demand (POD) edition of John Milton’s *Areopagitica* from an Internet Archive digitization of a University of California Libraries copy through to unedited digital and print version.<sup>16</sup> The Kindle’s

scale precludes thorough quality control. Amazon temporarily removes any reported low-quality editions, but this approach relies on free labor from customers who have purchased a copy of the text rather than Amazon manually or automatically identifying poor-quality digitization before publication. The automation of publication from archive to Kindle also produces books not intended for the audience. For example, on June 29, 2017, Project Gutenberg published the August 21, 1858, edition of *Stephen Branch's Alligator*, a regional newspaper focused on New York politics.<sup>17</sup> Prabhat Prakashan, a company specializing in Hindi literature, published a Kindle edition of the text, with clear attribution to Project Gutenberg, on September 16, 2017, for \$4.81.<sup>18</sup> No evidence indicates that anyone has bought the book, as it is available for free elsewhere. Amazon's commercial imperative runs counter to the ideology of other digitization projects—even for-profit organizations such as Early English Books Online (EEBO) and other UMI publications—as other initiatives value the archive as a unit over single texts. The values of Amazon's digitization efforts instead focus on offering already plentiful texts in a new format.

Amazon's digitization strategy differed from Google Books in two important ways. First, Google Books focused on quantity over quality, as most users would not see the underlying text and it formed the basis for a search engine, while Kindle users would purchase titles to read and expect a much higher quality. Second, Google Books' scope was necessarily wide to generate the most data, while Amazon needed to target titles that were feasible to sell. Eighteenth-century religious tracts and agricultural manuals enrich Google's search engine and draw users to the database but are unlikely to be commercially viable in bulking up an ebook store, especially once the expense of digitization is taken into consideration. In 2007, during the great push to convert materials for sale at the Kindle's launch, Levy reported that Amazon's cost for digitizing and correcting each book was roughly \$200 per title.<sup>19</sup>

In 2012, users could no longer download free editions of public domain texts but could access the same material through a Kindle Unlimited or Prime Reading subscription. Amazon's release of "Kindle in Motion" in 2017 attempted to distinguish Kindle public domain ebooks from other brands by adding animations and advanced typography and formatting to select public domain titles. Publishers could integrate animated GIFs into their books using the new feature. For example, Anna Sewell's *Black Beauty* is available as a Kindle in Motion edition through Two Lions, one of Amazon Publishing's imprints, to demonstrate the viability and power of the technology.<sup>20</sup> Pottermore is the only other publisher to produce Kindle in Motion editions to replicate the animated newspapers in *Harry Potter and*

*the Philosopher's Stone* and *Fantastic Beasts and Where to Find Them*. Amazon used public domain titles as a purchased proof of concept for the format rather than circulating for free. The strategy was unsuccessful, as no new Kindle in Motion titles were published after 2018. As discussed in chapter 4, Amazon's attempts to differentiate its format ecosystem have been met with publisher apathy without investment from the technology company, as publishers need to produce multiple copies at the lowest possible cost rather than optimize for each platform.

Beyond public domain titles, digitizing backlists of authors still protected by copyright represented a lucrative opportunity for those who held the rights. A legal distinction exists between copyright and rights. The former term refers to the exact expression, or the order of the words, while rights protect the idea in specific forms. For example, an author may transfer the rights to hardcover and paperback editions to different publishers. Adaptations, territories, and translations are all covered by separate rights, as well as different media, including audiobook and ebook. Each of these rights needs to be transferred in a contract. Sensing a gap in publishing contracts, Rosetta Books licensed ebook editions directly from authors such as Kurt Vonnegut and William Styron in 2001. Random House responded by suing the company in 2002.<sup>21</sup> The district court rejected these claims because the contracts did not specifically state "ebooks" but rather focused on physical rights.<sup>22</sup> The move was mirrored by Andrew Wylie, a prominent literary agent representing authors and estates including Chinua Achebe, Vladimir Nabokov, and J. M. Coetzee, who argued that since publishers had not negotiated for the rights for digital publication, he was able to publish these works through his own digital imprint, Odyssey Editions. Random House challenged Wylie in 2010, who settled the lawsuit. After the dispute, publishers began adding digital rights into contracts by default and quickly moved to ensure audio rights were included in contracts with the rise of Audible.<sup>23</sup>

Beyond recognized perennial sellers, the Kindle could amplify the discovery of obscurer titles and the backlists of authors to ensure a robust ebook marketplace without the costs of physically warehousing, printing, and keeping older titles in print. Cheap best sellers and free public domain ebooks attracted readers, but a healthy backlist ensured they stayed. Amazon introduced a "request this book" button on product pages not associated with a Kindle title to put pressure on publishers to convert popular print titles into an ebook. As many major authors were not instantly willing to offer their backlist on the Kindle Store, Amazon made announcements about major acquisitions, such as a deal with Penguin to publish Tom Clancy's backlist starting with *The Hunt for Red October* or the

arrival of Danielle Steel's back catalog, both in 2009.<sup>24</sup> As the platform matured, the announcements reduced in frequency with the expectation that authors' titles would be available by default.

The backlist is a central part of publishing business models. It is unusual for a reader to *only* consume new titles. The 80,000 titles available at the Kindle 1's launch were only a fraction of titles in print in November 2007: the Publishing Association estimated the number of new titles released in North America per year to be 150,000, almost double the titles available at the Kindle's launch.<sup>25</sup> The quantity was not exceptional, however, as competitors such as Microsoft Reader, a software package for personal digital assistants (PDAs), featured a library of sixty thousand titles in 2011. Ebook stores historically built catalogs on the new titles published as ebooks, and any titles already digitized from the public domain. This leaves the period between 1922 and the 1990s, which Rebecca Rosen has called the "missing twentieth century."<sup>26</sup> Out-of-print books protected by current copyright legislation can be difficult to find even from retailers such as Amazon. These sixty years represent a vibrant body of literature that many readers would like to revisit. The books remain in limbo, as the digitization costs would be borne entirely by publishers. Investment in the back catalog only pays off with titles that sell. Despite the risk, ebooks are useful for testing the potential audience for a backlist title before investing in a full print run.

The digitization of the missing twentieth century is understandably patchy. Google Books and HathiTrust digitized many twentieth-century publications without permission from the rights holders. Copyright challenges prevent the data from being shared. Beyond intellectual property issues, the cost of digitization creates a barrier for many. Publishers including the MIT Press often rely on external funding from bodies such as Arcadia and the Internet Archive to incentivize digitization projects.<sup>27</sup> To determine the extent of digitization for the Kindle, I compared titles listed in the British National Bibliography as published in 1989 (30,490 books) with the Amazon store. These books' availability on the Kindle and, more broadly, the Amazon store reveals stalled progress in digitizing a cross section of the backlist. The year 1989 is significant in the history of digital books, as it marked the launch of Project Gutenberg's first ebook. The year is recent enough for books to potentially be topical and remain in print, yet distant enough that publishers were still shifting to digital workflows, so that most new publications would involve a hybrid of traditional and digital techniques. Amazon uploaded the data from the 1995 edition of *Books in Print* for its initial catalog, which was kept separate from *Books Out of Print 1995*, which contained 1989 titles that were out of print six years



later. The product pages therefore indicate the proportion of titles still in circulation on the secondhand market and how many of those titles have appeared in a digitized form as a Kindle edition.

Table 5.1 breaks down the availability of titles published in 1989, where fewer than 10 percent of all titles have received a Kindle edition, with only 1 percent present in the Kindle Unlimited scheme, indicating that the service's exclusivity clause discourages publishers from sharing their backlist in the scheme. Amazon stocks more books in physical inventory than are available through the Kindle, and this is far surpassed by the availability of the books by third-party retailers. The third-party retail economy ensures the continual circulation of these out-of-print titles. Used copies of books from libraries are sold through Amazon, but if titles are no longer in circulation, we have no clear evidence the books may sell.

The overall volume of “Look Inside the Book” titles outweighs the number available on the Kindle, which automatically includes the feature on product pages. Almost two thousand titles—the equivalent of 80 percent of the volume of titles available for the Kindle—are available to browse online but are only sold in print via either Amazon or third-party retailers. These titles are digitized and shared with Amazon but have not been made available to sell on the platform, indicating there are social rather than technical issues for excluding the titles from the platform. Two main reasons exist for this gap. First, as discussed in chapter 1, the “Look Inside” function predates the Kindle and was used as an incentive for marketing,

**Table 5.1** Digitization progress for books published in 1989\*

Availability	Frequency
Kindle	2,697
Kindle Unlimited	201
First-party print	3,332
First-party print with Look Inside	1,142
Third-party print	21,974
Third-party print with Look Inside	642
404 error	60
No available copies	5,574
<b>Total</b>	<b>30,940</b>

\*The sixty 404 errors come from books published by small independent publishers and individuals as well as ISBNs that were reused for several books by large academic publishers such as Elsevier that are not easily compatible with the ASIN infrastructure.

and publishers may not have granted Amazon the additional permissions to convert the book to a Kindle edition. “Look Inside” reproduces a facsimile rather than requiring the costlier step of extracting the text via optical character recognition to create a publishable ebook. Second, the genre of many books available exclusively via “Look Inside” reveals variations in digital distribution, as large academic publishers such as Springer Nature and Elsevier used the marketing potential of “Look Inside” but elected to not sell the books via the Kindle. Scholarly publishers invested in digital workflow technologies earlier than other publishers and gain more benefits from digitizing the rest of their back catalog owing to the additional merits of print-on-demand business models for the scale of print runs in academic publishing. The sector set up alternate infrastructures for disseminating ebooks on platforms such as ScienceDirect. As the library subscription model is the primary revenue stream for academic publishers, they rely less on the availability of titles for the Kindle.

The historical and geographical specificity of publishing also affects the availability of ebooks from 1989. While readers still consume classic works, most publications remain forgotten. The Cold War, the HIV/AIDS crisis, and futurist speculations on the 1990s were popular topics in 1989 publishing. Some of these titles may remain appealing, but the majority are unlikely to generate sufficient sales beyond curiosity to be profitable enough to digitize. The long tail of publishing uncovered by analyzing a slice of all books published in a single year also reveals a body of local-history titles that would not have a large-enough scale of interest to prompt a digitization campaign.

In a talk at Ebookcraft 2016, Teresa Elsey, then digital managing editor at Houghton Mifflin Harcourt, contextualized the lack of progress with the backlist by demonstrating the pressures of maintaining a backlist that remains perpetually “in print.” Elsey calculated her group’s workload for 2015—a total of 1,400 ebooks—of which only 20 percent were new titles. The remaining titles were books that had never received an ebook edition (10 percent) and new editions of preexisting ebooks (70 percent).<sup>28</sup> As more titles are introduced as ebooks, the burden of updating becomes a large proportion of the workload, leaving the older titles less likely to be digitized. Elsey attributes this constant churn to two factors: (1) the initial digitization gold rush, and (2) the shifting demands of ebook production and the anachronistic connection between ebook titles and reading systems. The digitization gold rush in the mid-2000s, facilitated by Google and Amazon, emphasized quantity over quality, and consequently an abundance of OCR errors, uncorrected print-based hyphenations, and other conversion artifacts remained in the process; in today’s mature

ebook marketplace, these errors are noticed by retailers and readers and need to be corrected. Elsey also notes the importance of the legal issues regarding permissions for digital editions that have emerged with a more substantial effort to digitize the backlist. The shifting demands of ebook production affect all backlist titles, including those born digital. Older titles need maintenance according to updated standards, and the conventions of ebook design shift, so even working titles may require updating to reflect the market. This is exacerbated by the anachronistic connection between titles and reading systems, which show no sign of aging, unlike their print counterparts.<sup>29</sup> Maintaining a backlist with the pressures of updatability remains an open challenge a decade after the Kindle's launch, decreasing the resources available for publishers to digitize more of their backlists.

### **The Emergence of Digital-Only Publishers**

The now-infamous *Fifty Shades of Grey* trilogy challenged orthodox conceptions of publishing success made possible through developments with ebooks.<sup>30</sup> E. L. James, the pen name of Erika Mitchell, presented a mythical “rags-to-riches” narrative highlighting the potency of self-publishing. The novel started as fan fiction of the *Twilight Saga* posted on fanfiction.net, published in weekly installments for a rapt audience. Once James found a large audience, she published a revised version of the fan fiction as *Fifty Shades of Grey*, using print on demand before Vintage Books acquired the rights. James initially found success without the support of a traditional publisher, but her biography reveals that this is not a straightforward narrative. Before the publication of *Fifty Shades*, James worked in television, and her husband, Niall Leonard, remains a screenwriter. James had appropriate connections to media industries and understood the nuances of contemporary publishing when choosing to self-publish the work.

Given the book's origins as fan fiction, it could not be published in its original form. James's media-savvy marketing through fanfiction.net built an audience that an author taking the traditional approach would struggle to replicate. The blurring of the lines between best seller and self-published work that *Fifty Shades of Grey* pioneered reveals the blueprint for entrepreneurial authors willing to start with the Kindle. The ability to read “mommy porn” inconspicuously was immaterial compared to its use of self-publishing, as the success of *Fifty Shades* depended equally on the print edition and its appearance in various gendered locations, including supermarkets. For example, in Norway, 96 percent of 177,000 total sales

for the *Fifty Shades* trilogy were print rather than digital.<sup>31</sup> *Fifty Shades* sold over eleven million paperback copies in the United Kingdom in its first five years, according to Nielsen BookScan, further solidifying the series' best-seller status. The third title, *Fifty Shades Freed*, was the weakest seller but still sold twice as many copies as the fourth most purchased book from 2012 to 2017, *Harry Potter and the Cursed Child*. *Fifty Shades* was a phenomenon across formats that revealed the scale of the ebook market during its ascent. Regardless of the authenticity of James's underdog narrative, *Fifty Shades* demonstrated the marketing potential of fan fiction forums and self-published books to build loyalty before embarking on a more traditional publishing model.

The rise of self-published authors reverses twentieth-century trends but fits into more historical trajectories of publishing. The publisher has not always been separated from the author; for example, Samuel Richardson was the author, publisher, and printer of *Clarissa* in the 1740s. It was only with the professionalization of the printing press and publishers that Richardson's do-it-yourself mentality became abnormal. Self-publishing was tarnished as vanity publishing, since if a traditional publisher deemed a book unworthy, it was unlikely to be worth the money. Vanity publishing required authors to subsidize the cost of production and place advertisements at the end of the book to ensure a financial return for the printers.<sup>32</sup> Predatory vanity publishers, particularly in academic publishing, generated a stigma around self-publishing that presented a hurdle for the legitimacy of services such as Kindle Direct Publishing.

Kindle 1 promotional material heralded its extensive range of *New York Times* Best Sellers and backlist titles, but the launch of Digital Text Platform (the original name of Kindle Direct Publishing) marked a more substantial transformation. Amazon did not want to reveal its self-publishing ebook platform before the launch of the hardware, but the Digital Text Platform page launched on the same day as the Kindle and encouraged users to join up with Amazon. The splash page was accompanied by an FAQ that sold Amazon's ability to act as an intermediary in place of the publisher because "the Kindle Store provides the business, marketing, and vendor support you need to maximize your sales. Titles published on the Kindle Store are eligible for automated merchandising and personalization, Search Inside, Amazon.com's excellent Customer Service and order fulfillment services, and much more."<sup>33</sup> The FAQ, though not publicly released until the November launch date, was written in May 2007. KDP represented a strong backup plan for Amazon if print publishers were unwilling to publish digital publications.

Despite the term “direct publishing,” authors using Kindle Direct Publishing benefited from association with the Amazon brand. Rather than releasing a MOBI file to a web page and hoping for the best, self-published titles were boasted by Amazon’s full-service infrastructure and discoverability platform. While traditionally published and self-published works largely remained separate, authors could leverage the benefits of Amazon as de facto publisher in the process. Kindle Direct Publishing was not just another place to sell a few ebooks; it offered a new way for authors to make a living without relying on traditional publishing structures. The company’s role as publisher is clear from how it shapes pricing: while authors can charge whatever they like, in 2015, authors only enjoyed a 70 percent royalty if they priced their title between \$3 and \$10.<sup>34</sup> Incentivizing lower prices encouraged the production of material deemed viable at that cost. KDP is not self-sufficient, however, as it relies on the prestige of legacy publishing to ensure that readers use the platform. The size of the market attracts users and authors to the platform, but without the prestige of a publisher’s imprint, the Kindle platform would compete with start-ups like Wattpad and Archive of Our Own that offer similar services for free. These two divergent markets have been the subject of intense trade scrutiny, as official figures from the Publishers Association noted a 17 percent drop in sales of consumer ebooks in 2016, resulting in joyous articles about the downfall of ebooks.<sup>35</sup> The figures exclude two large segments of the market: digital-only publications and the entirety of Kindle Direct Publishing.

Early success stories shaped Kindle Direct Publishing’s impact. John Locke capitalized on his fame as the first self-published author to sell over a million books. Locke was fortunate to be the first author to hold the top two slots in most-purchased Kindle titles.<sup>36</sup> His book covers reference the large volume of titles that have been sold, and a sense of bravado permeates his author brand. Locke signed a distribution deal with Simon & Schuster in mid-2011, allowing him to continue self-publishing new works.<sup>37</sup> This unusual deal marked a transition point where self-publishing was now seen as a viable career option, shedding its historical association with vanity publishing. Locke was later implicated in paying reviewers to secure a five-star rating for his titles and furthering the tactic by encouraging reviewers to purchase an Amazon-verified copy of the book to both increase his sales ranking and game Amazon’s review algorithm.<sup>38</sup> The investigation raised important questions about the ethics and sustainability of self-published best sellers, which would often have to be priced at less than a dollar to be considered, compared to the higher price points

expected from print-oriented publishers. After this negative publicity, and after the initial lucrative flurry of sales diminished, Locke stopped publishing new material in 2012.

The Locke model for self-publishing relied on the early Kindle marketplace, and as it has matured, so has the role of self-published and digital-only publications. The arrival of Kindle Unlimited, and more recently Prime Reading, is bolstered by the presence of these types of publication. Select popular legacy publications like the *Hunger Games* trilogy and *Harry Potter* series pepper a catalog of over a million titles. The tactic worked for Suzanne Collins, as the *Hunger Games* books were revealed as three of the top five most purchased fiction titles in the Kindle's first decade, alongside the first two *Fifty Shades* titles.<sup>39</sup> Collins outperforms Rowling because the latter entered the platform later and sold books via a third party. Licensing conditions for Kindle Unlimited require publishers to publish exclusively via the Kindle platform, optimizing the scheme for Kindle-exclusive titles available via KDP while remaining a barrier for traditional publishers.<sup>40</sup> The subscription service offered a revenue kitty to be shared according to volume of pages read. In August 2017, the KDP Select Global Fund was \$19.4 million, although the payout per page was 0.4 cents, indicating that subscribers viewed more than 4.6 billion "pages" per month.<sup>41</sup> In this competitive attention economy, authors moved from Locke's scheme of paying for reviews to using underhanded tactics to encourage readers to skip to the end of the book or otherwise increase the visibility of a title.

Locke's early exploitation of Kindle Direct Publishing's fragility has been usurped by more flagrant abuses in recent years, including accusations of money laundering through selling extremely expensive fake books using authenticated authors' credentials.<sup>42</sup> David Gaughran, an author and advocate for self-publishing, has documented a litany of algorithmic gaming practices and errant punishments to authors who follow the rules to crack down on unscrupulous behavior. The Kindle Unlimited royalty pot has led to a common fake book scam, where content is stolen and rendered in large 2,500-page volumes, offered for free, and then added to Kindle Unlimited once its visibility is high to capture a larger proportion of the KDP royalties that month.<sup>43</sup> Authors and publishers of original content have also sunk to the same lows of click farming, or paying for third parties to download, "read," and review a book to increase its visibility.<sup>44</sup> Amazon's response has been to punish these unusual behaviors, but with the unintended side effect of harming authors who use third-party services such as BookBub to advertise their books.<sup>45</sup>

The benefits of self-publishing for groups underrepresented by print-oriented publishers have been overshadowed by automatically generated

content that floods the Kindle Store.<sup>46</sup> Despite various schemes to undermine the credibility of self-publishing via Amazon, several authors have set an example of the new careers available in the digital age. Amanda Hocking, the second self-published author on the Kindle Millions Club, was one of the forerunners of the second wave of self-publishers. Hocking released her first book, *My Blood Approves*, in March 2010, and by the time she had sold a million books, she had published two series: *My Blood Approves* and *The Hollows*. The second book in *The Hollows* series was published a day before her induction into the Kindle Million Club, indicating a surge of sales related to this title. St. Martin's Press signed Hocking in 2011, but before that point, she was selling an average of nine thousand books a day.<sup>47</sup> Hocking's success demonstrates that self-publishing ebooks could lead to sustained print sales, as her most recent book series, *The Kanin Chronicles*, was published in 2015 by St. Martin's Press, four years after the start of the relationship. Self-publishing platforms such as Kindle Direct Publishing allowed authors such as E. L. James and Amanda Hocking to launch their careers, but both authors chose a more traditional route once they had established track records. These success stories have encouraged others to self-publish, with many authors electing to stick with self-publishing because of the autonomy and royalties unmatched by legacy publishers.

The rise of digital-only publishers like Canelo, a team who drew on experience from print publishers including HarperCollins and Bloomsbury, shows a sustainable approach to the development of separate print and digital reading audiences. The ebook marketplace has matured around the expectations of value from a consumer perspective. While readers may spend up to \$25 or £20 on a newly released hardback book, the equivalent ebook's value proposition is far trickier. The perception that ebooks should be cheaper has led to more agile publishing models for digital-only titles, where publishers and self-published authors use newly created spaces within the ebook market. For example, ebooks can be much shorter than print equivalents, which makes it viable to charge as little as a dollar for a single book. Canelo's mission statement lists what it considers to be viable digital markets: crime, commercial women's fiction, action fiction, and "what we call 'geeky crossover,' such as horror, or historical fantasy."<sup>48</sup> The company focuses on genres it believes will thrive at a low and affordable price. Canelo expanded its offerings by opening Abandoned Bookshop as an imprint in 2016, with a mission statement to "uncover the best books we've forgotten about, lost sight of, or never even knew existed in the first place."<sup>49</sup> Focusing on recovering forgotten titles— orphaned works, titles out of copyright, or simply books that were not paid attention by their

original publisher—approaches the backlist as an opportunity rather than a risk. Abandoned Bookshop leverages the credentials of publishers as curators to differentiate the venture from companies that have loaded portions of their backlist to the Kindle Store without a clear focus, and the mass of self-published work that defines digital-only publishing today.

Amazon's role in the distinction between self-publishing and digital only is complex. Since Amazon provides many services associated with the publishing process, can Kindle Direct Publishing be called self-publishing rather than traditional publishing, albeit with little editorial intervention? Amazon published 1.6 million titles through Kindle Direct Publishing in 2016, which would make the company the largest publishing company in the world.<sup>50</sup> It has run various schemes that fit into legacy publishing models, but with a digital twist. In 1999, “the company quietly leased the rights to a defunct imprint called Weathervane, and repackaged . . . creatures from the black lagoon of the remainder table.”<sup>51</sup> The project was short-lived and erased from Amazon's institutional memory. The company returned to the idea while gearing up for the launch of the Kindle with the 2005 launch of Amazon Shorts, later retitled Kindle Singles. The program offered authors the ability to publish “short-form work (2,000–10,000 words, fiction or nonfiction)” that would be “available to customers via PDF, HTML, plain-text e-mail, and they are stored in customers' Digital Lockers forever.”<sup>52</sup> The unfortunate mention of permanence in the marketing of an obsolete product demonstrates the precarity of the Kindle within Amazon's ecosystem. The company invested in the scheme in the signature Amazon style of commissioning exclusive content for the launch of the Kindle 2: Danielle Steel penned an exclusive autobiographical essay, “Candy for the Soul,” for the launch of Amazon Shorts. In a 2012 editorial for the *New York Times*, Dwight Garner praised the series as “probably the best reason to buy an e-reader in the first place” as the home for “works of long-form journalism that seek out that sweet spot between magazine articles and hardcover books.”<sup>53</sup> Enthusiasm for the format led to 1,248 publications in the series, but production slowed in recent years from a peak of 242 titles in 2015 to just 109 in 2017.

Since the Kindle's launch, Amazon launched a further fourteen imprints, ranging from Waterfall Press, a specialist Christian publisher, to Amazon Crossing, which specializes in translations. The imprints that focus on popular trade fiction, including most prominently Thomas & Mercer, a mystery and thriller imprint (911 titles since 2011), and Montlake Romance (1,378 titles since 2011), have had the greatest impact on publishing. Amazon publishes print copies of these texts as part of the promotion, although other retailers have hesitated to stock a book published by their



fiercest competitor.<sup>54</sup> Through all these imprints, Amazon has published eight hundred books per year since 2009, making the company one of the twenty largest publishers in the United States.

### Coda: Amazon Charts

May 2017 saw the launch of Amazon Charts, a weekly update on the most read and purchased books across the US Amazon web store, the company's latest attempt to disrupt trade publishing orthodoxies. The *Washington Post*—owned by Jeff Bezos, albeit independently of Amazon—included the data in its best-seller lists in January 2018 to provide the first list to account for the largest ebook retailer.<sup>55</sup> Amazon Charts data did not come with any figures but offered an inclusive definition of “book,” which encompassed print, ebooks, and Audible. The latter two categories provided observable data for the “Most Read” charts. Kindle Unlimited books were counted “once a customer has read a certain percentage—roughly the length of a free reading sample,”<sup>56</sup> offering a surprising degree of transparency regarding methodology compared to the *New York Times* Best Sellers. The list also includes other features, such as tidbits about the authors' agents, that are unlikely to interest nonindustry members, as well as facts about the books based on the data, including tagging books as “unputdownable” if readers had generally finished the book in a single session, an honor that Sarah J. Maas's *A Court of Wings and Ruins* received in the first week.<sup>57</sup>

The separation of “Most Read” and “Most Sold” reveals the limitations of focusing on sales, which might not map over to what readers consume. I analyzed Amazon Charts' first eighteen weeks between May and September 2017 to show how the two categories work in parallel. Unsurprisingly, both reading and purchasing habits are strongly tied to film and television adaptations: Margaret Atwood's *The Handmaid's Tale* was omnipresent in both lists due to the television show's run from April to June and its availability on Kindle Unlimited. Negative reviews of a film have a chilling effect on Amazon users' purchasing and reading habits. For example, Stephen King's *The Gunslinger* appeared on the “Most Sold” list between May and July 2017 in anticipation of the film's launch on August 20. The book appeared on the “Most Read” lists for four weeks between July 30 and its cinematic release, when interest immediately dissipated. The influence of Kindle Unlimited is stronger for fiction than nonfiction. For example, the *Harry Potter* titles frequently appear in the “Most Read” section, but as the series is exclusively available via Kindle Unlimited, this does not register as purchasing the title. The data demonstrate a distinct split between hype

cycles and marketing campaigns fueling purchases and what readers are actually interested in.

Amazon Charts provide a useful counterpoint to the *New York Times* Best Sellers list, a barometer of taste and frequent reference point for Amazon during the development of the Kindle. Thirty-nine percent of fiction and 43 percent of nonfiction that appeared on the *New York Times* Best Sellers list also appeared on Amazon Charts. More books appear on both the *Times* list and Amazon's "Most Read" books (47 percent of fiction titles, 51 percent for nonfiction). Readers are drawn to the same big books—Atwood's *The Handmaid's Tale*, Neil deGrasse Tyson's *Astrophysics for People in a Hurry*, Trevor Noah's *Born a Crime*—but differences emerge in the titles that hover farther down the charts. Some of these variations may be a matter of sampling. For example, Stephen Covey's *The 7 Habits of Highly Effective People* would be classified as a perennial seller by the *New York Times* and would be ignored even if it outsold other titles. Elsewhere the differences show how agency pricing has warped the market. Stuart Wood and Parnell Hall's *Barely Legal* debuted at number 5 on the *New York Times* Best Sellers list on August 13, 2017, but it did not appear on Amazon Charts. In September 2017, the Kindle edition was priced at \$17.54, a mere fifteen cents cheaper than the hardcover. Since consumers value digital goods less than the equivalent physical artifact, the lack of discount has boosted print sales at the expense of ebook growth in mainstream publishing.<sup>58</sup>

Amazon released the best-selling Kindle title for each of the first ten years to celebrate its first decade (table 5.2).<sup>59</sup> The international launch of the Kindle in 2009 leans toward a convergence with the print marketplace, but the first two years are more intriguing. The 2008 best seller, *The Complete User's Guide to the Amazing Amazon Kindle*, outsold Khaled Hosseini's *A Thousand Splendid Suns* and Patricia Cornwell's *Book of the Dead* as users wished to learn more about the new platform. The inclusion of Ken Follett's *The Pillars of the Earth*, a novel first published in 1989, is more intriguing. Follett's sequel, *World without End*, was published in October 2007 and was one of the featured \$9.99 *New York Times* Best Sellers on the first-generation Kindle's launch page. The prequel outsold the original, as it was both cheaper (\$6.39) and longer (1,000 pages), demonstrating the volatility of ebook pricing from the start of the Kindle's life span.<sup>60</sup> The ability to read a dense novel from the backlist for cheap without carrying it around marked a landmark moment for the early adoption of the device as users began to see its advantages over print.

The Kindle's first decade has seen a wave of experimentation and the divergence of print and digital marketplaces. This shift has not necessarily

**Table 5.2** Kindle year-by-year best sellers\*

Year	Kindle best seller
2007	Ken Follett, <i>The Pillars of the Earth</i> (1989)
2008	Stephen Windwalker, <i>The Complete User's Guide to the Amazing Amazon Kindle</i> (2008)
2009	Dan Brown, <i>The Lost Symbol</i> (2009)
2010	Stieg Larsson, <i>The Girl with the Dragon Tattoo</i> (2008)
2011	Kathryn Stockett, <i>The Help</i> (2009)
2012	E. L. James, <i>Fifty Shades of Grey</i> (2011)
2013	Dan Brown, <i>Inferno</i> (2013)
2014	John Green, <i>The Fault in Our Stars</i> (2012)
2015	Paula Hawkins, <i>The Girl on the Train</i> (2015)
2016	Paula Hawkins, <i>The Girl on the Train</i> (2015)
2017	Margaret Atwood, <i>The Handmaid's Tale</i> (1985)

\*Wheat, "10 Years, 10 Books."

democratized contemporary publishing, as predicted by early advocates of self-publishing. As indicated by the presence of the legacy publishing heavyweights in the "Most Read" section of Amazon Charts, the industry still depends on hype cycles, perennially popular titles, and best sellers. The initial promise of digitization and the backlist did not materialize into substantial new business for legacy publishers. While legacy publishers continue to engage with digital publishing, albeit with the caveats of agency pricing or with other concessions from retailers such as Amazon, the self-publishing and digital-only models will continue to thrive; but if the scaffolding from legacy media companies is removed, there is less evidence to demonstrate that this community is self-sustainable, and much of the cultural record may disappear with the planned obsolescence cycles of digital media. The ebook marketplace exists in a precarious state, still reliant on print-oriented publishers that resist committing to digital publishing. Amazon's infrastructure is essential to maintaining a strong presence for ebooks, but the lack of trust between publishers and Amazon continues to erode public interest in ebooks. The Kindle Store requires both continued investment from print-oriented publishers and the encouragement of digital publishing start-ups to become sustainable.



This is a section of [doi:10.7551/mitpress/11985.001.0001](https://doi.org/10.7551/mitpress/11985.001.0001)

# Four Shades of Gray

## The Amazon Kindle Platform

By: Simon Peter Rowberry

### Citation:

*Four Shades of Gray: The Amazon Kindle Platform*

By: Simon Peter Rowberry

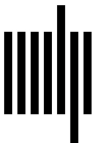
DOI: 10.7551/mitpress/11985.001.0001

ISBN (electronic): 9780262369114

Publisher: The MIT Press

Published: 2022

The open access edition of this book was made possible by generous funding and support from MIT Press Direct to Open



The MIT Press

© 2022 Simon Peter Rowberry

All rights reserved. No part of this book may be reproduced in any form by any electronic or mechanical means (including photocopying, recording, or information storage and retrieval) without permission in writing from the publisher.

The MIT Press would like to thank the anonymous peer reviewers who provided comments on drafts of this book. The generous work of academic experts is essential for establishing the authority and quality of our publications. We acknowledge with gratitude the contributions of these otherwise uncredited readers.

This book was set in Filosofia OT by Jen Jackowitz. Printed and bound in the United States of America.

Library of Congress Cataloging-in-Publication Data

Names: Rowberry, Simon Peter, author.

Title: Four shades of gray : the Amazon kindle platform / Simon Peter Rowberry.

Description: Cambridge, Massachusetts : The MIT Press, [2022] | Series:

Platform studies | Includes bibliographical references and index.

Identifiers: LCCN 2021013279 | ISBN 9780262543507 (paperback)

Subjects: LCSH: Kindle (Electronic book reader) | Electronic book readers.

| Electronic books.

Classification: LCC Z286.E43 R689 2022 | DDC 004.1675—dc23

LC record available at <https://lcn.loc.gov/2021013279>

10 9 8 7 6 5 4 3 2 1