

From publishers to self-publishing: Disruptive effects in the book industry

Morten Hviid^{*}, Sofia Izquierdo-Sanchez[†] and Sabine Jacques[‡]

Abstract

This paper explores the structure of the book publishing industry post-digitalisation, analysing the choices of the publishers and authors. The introduction of successful e-book readers has belatedly given digitalisation the characteristics of a disruptive technology by making self-publishing a serious option for authors. This has been supported by the entry of new types of intermediaries and the strengthening of others. These changes have reduced the general requirements for an author to get a book self-published. As a result, a larger share of the surplus from the book industry is likely to go to authors, explaining the significant increase in the supply of books. The potential over-supply of books has created a new problem by increasing competition and making consumer searches more difficult. We argue that digitalisation has shifted the potential for market failure from an inadequate supply of books to asymmetric information about quality. It remains to be seen whether the market will provide appropriate intermediaries to solve the asymmetric information problem.

Keywords: Book industry; Self-publishing; Traditional publishing; Retailers.

JEL classification: D22; D4; L82

^{*} Centre for Competition Policy and School of Law, University of East Anglia, Norwich, NR4 7TJ.

[†] Corresponding author: Department of Accountancy, Finance, and Economics and Northern Productivity Hub. University of Huddersfield, HD1 3DH. E-mail: s.izquierdo-sanchez@hud.ac.uk

[‡] Centre for Competition Policy and School of Law, University of East Anglia, Norwich, NR4 7TJ.

1. Introduction

For decades, if not centuries, traditional publishers have been the gate-keepers of the book industry. They have provided an essential link between the author and the audience because they possessed superior data and experience, both in terms of assessing manuscripts and producing, distributing and promoting the resulting books (Greco, 1945 and revisited editions). The most recent challenge to this position has been the arrival of e-readers, enabling thousands of e-books to be carried and read on a single device. While electronic versions of books, now referred to as ‘e-books’,¹ have been available since at least the 1970s [see Gilbert (2015) and Waldfoegel and Reimers (2015)], it required the introduction of a dedicated reading device for the sales and use of e-books to truly take off. The launch of the first commercially successful e-reader, the Amazon Kindle in 2007, triggered a dramatic change in the book industry.² By the end of 2013, the share of e-book sales in the US reached almost a quarter of all book sales, [see Gilbert (2015, page 166)] while for the UK, the official market share of e-books reached 21% of the total retail market by value.³

Prior to digitalisation, it would have been prohibitively expensive for most authors to bypass publishers through self-publishing.⁴ However, this has changed with the arrival of the e-book and, in particular, e-book readers. Although traditional publishers still have a comparative advantage in terms of better data and funds to pay advances, self-publishing e-books has benefits in terms of speed as well as offering the author greater freedom and control over the final product and, possibly, even the pricing of their book. The emergence of various internet-based services has also made self-publishing a more feasible choice for authors. Firstly,

¹ The terminology is not fully settled as yet and others use the variations ‘ebook’ or ‘eBook’.

² While digitalisation of books started in 1971 (with project Gutenberg) or possibly even earlier and e-readers also emerged relatively early, the first commercially successful e-book reader was the Amazon Kindle.

³ The US and UK are the two largest markets in terms of e-book sales [see: <https://www.thebookseller.com/news/self-published-titles-22-e-book-market-325152>]. What does emerge from existing industry data is a clear indication that self-publishing is a potential game-changer. Bowker, using data on books with ISBN’s find that the number of self-published ISBN rose from 247,210 in 2011 to 786,935 in 2016, [see <http://www.bowker.com/news/2017/Self-Publishing-ISBNs-Climbed-8-Between-2015-2016.html>]. As a consequence, we need to rethink the regulatory structure surrounding the industry and to evaluate industry behaviour and lobbying in light of this. While US sales of e-books were reported to have fallen in 2015 [see: <https://www.theguardian.com/books/2017/mar/14/ebook-sales-continue-to-fall-nielsen-survey-uk-book-sales>], this has been hotly disputed. Research undertaken by *Author Earnings* (<http://authorearnings.com>) finds that an increasing number of e-books do not have an ISBN number and such books would not be accounted for in the traditional statistics. Hence, the data on which the claim made about declining sales was based under-represented e-books.

⁴ However consider the counter-example of John James Audubon whose first book, “The Birds of America” was self-published and essentially crowd-funded through subscription. Creative and fortunate authors may always have been able to get around the gate-keeping publishers.

powerful internet-based retailers, such as Amazon, are willing to deal with self-publishing authors and to provide them with direct access to consumers. Secondly, over the last 10 years, a number of intermediaries have entered the market, offering to carry out several of the steps involved in converting a draft manuscript into an e-book (and even a printed book) or to provide the software and support to enable the author to carry out these tasks themselves. The effect has taken time to emerge but has recently become significant.⁵

The aim of this paper is two-fold (1) to analyse the publishers' best response towards the self-publishing market; and (2) to examine the option of self-publishing for authors and to assess whether it represents a realistic alternative to traditional publishing using conventional intermediaries (such as publishers). In particular, we seek to understand what features make it a viable alternative, and whether there is a danger that the option of self-publishing may be curtailed in the future. Secondly, we investigate whether the ability to self-publish affects the bargaining power of authors and, hence, the allocation of revenue among those involved in creating and producing the final product, or whether it mainly allows authors more control over the final product. The focus is predominantly on the UK with occasional comparisons with other markets, particularly but not exclusively the US. The motivation for looking at the UK comes in part from an assessment made in Thompson (2012, chapter 8) that the UK is the "wild west" of publishing. As the "wild west" is where rules are made, challenged or broken it serves as an interesting laboratory for the study of the effects of a disruptive digital innovation.

The paper is organised as follows. The following section describes the e-book publishing industry pre-digitalisation and then identifies what has changed through digitalisation the emergence of disintermediating entrants. Section three focuses on the legal aspects, and unpicks the relationship between the book publishing industry and the law in the digital context. The primary focus is to identify copyright issues which arise in relation to the self-publication of e-books. Section four explores an author's choice first between doing (self-publishing) and delegating (using a publisher) and then between what the author does and what services are bought in. Section five considers how the publishers might optimally respond to the increased ability and willingness of authors to self-publish, either by scrutinising manuscripts better or by waiting and seeing how the market views the new offerings. Section six discusses the

⁵ Thompson (2012), a tour-de-force of the history and current reality of book publishing, hardly devote any time to self-publishing – it only gets one mention, on page 154. The development in the five or so years since the publication of this book is very marked and largely justifies the more optimistic views on developments. Thompson argues that the UK publishes more books per capita than any other country, furthermore they feature a small number of very large corporations controlling a significant share of the market

problem arising from the long tail. Section seven summarises the conclusions and points to likely future developments.

2. The structure of the market pre-digitalisation and subsequent changes

Pre-digitalisation, the vertical structure of the book market could be illustrated as in Figure 1 below.

[Figure 1 around here]

The simplified figure includes the key actors but, in addition to this simple model, many authors hire agents to represent their interests and many publishers use independent wholesalers or distributors to supply the retailers (Thompson, 2012).

2.1 Publishers and publishing

As illustrated by Table 1 below, there are a large number of publishers in the UK publishing industry. Moreover, while the number of UK publishers has fluctuated over the last 14 years, it has done so around an average of approximately 2200 with a short-lived spike in 2008.⁶

[Table 1 around here]

At the same time, as shown in Table 2 below, the level of concentration in UK book publishing has also remained fairly constant. Analysis of the four-firm concentration ratio - measuring the market share of the four largest publishers - indicates that these publishers have retained almost half of the market.⁷

[Table 2 around here]

Moreover, data from the BA Reports Library indicates that, since 2009, the UK market share of the fifth largest publisher has been below 4%. Thus, there is a dramatic drop-off in size outside the top four publishers. The numbers in Table 2 do, however, mask a change in concentration. The market share of the largest firm, Bertelsmann, jumped from 14.8% in 2012 to 24.1% in 2013 following a merger with another of the largest firms.⁸ Although less frequent

⁶ These numbers have to be viewed with some caution. We know that there is considerable merger and acquisition activity in this sector. It is not clear whether at acquisition the old brand or VAT number ceases to exist. Thus, the number of VAT registered publishers may not correspond to the actual independent decision makers in this industry.

⁷ The numbers in the table may overstate the actual market share of the large publishers because of an increase in e-books without an ISBN which would not be accounted for in the official data (*Author Earnings*: <http://authorearnings.com/report/november-2015-the-uk-report-author-earnings-on-amazon-co-uk/>)

⁸ The £2.4bn merger between Penguin Books and Random House, creating the largest book publisher in the industry [see: <https://www.telegraph.co.uk/finance/newsbysector/mediatechnologyandtelecoms/media/9639925/Penguin-and-Random-House-to-merge-in-2.4bn-deal.html>].

than in other creative industries, mergers and acquisitions do occur in the book publishing industry but without creating significant changes in concentration among the Top 4 or 5 firms. Therefore, the overall picture in the UK is one of a very small number of large publishers each holding a large portfolio of book titles and enjoying large sales. At the same time there is a long tail comprising smaller, often specialist, publishers and an increasingly significant number of self-publishing indie authors.⁹ A similar picture characterises other markets; globally the English language book market is dominated by five large publishers, often referred to as the Big Five.¹⁰ The market shares may understate the actual degree of competition among publishers (Latcovich and Smith,2001). The US Second Circuit Court decision in the Apple e-book competition litigation noted that key players in the publishing sector had no qualms about holding regular meetings to discuss common challenges they faced.¹¹ This begs the question of exactly how competitive this industry is and has been. The stability of the market structure is interesting in itself, because it suggests that, up until now at least, the traditional publishing ‘powerhouses’ have been able to maintain their position despite the very technological innovations which might be expected to challenge their position.

2.2 Book retailing

The decades leading up to the introduction of the e-book reader saw the creation of powerful retailers with the ability to challenge the large publishing houses.¹² While the challenge initially came from the creation of bookshop chains, the eventual challenge arose from the creation of online booksellers and, in particular, the launch of the Amazon Kindle Store.

The e-book market is dominated by one retailer, Amazon. *Author Earnings* estimates Amazon’s share of US e-book sales at 74%, with the majority of the remaining share being split among four other retailers: Apple’s iBooks Store (11%), Barnes & Noble’s Nook Store (8%), Kobo’s US bookstore (3%), and Google Play Books [2%], so thereby only 2% for sales via other avenues. The US is the largest market for e-books, with the UK in second place. *Author Earnings* has recently extended its empirical work to cover the UK market.¹³ This

⁹ This feature is not unique to book publishing. Waldfogel and Reimers (2015) mentions a similar structure in the music industry.

¹⁰ This used to be the Big-Six: Hachette, Harper Collins, Macmillan, Penguin, Random House and Simon & Schuster. Following the merger of Penguin and Random House, the current Big-Five are: Bertelsmann (Penguin Random House/Transworld), Hachette Livre UK (Headline/Hodder/Little Brown/Orion), News Corporation (HarperCollins), Holtzbrinck (Pan Macmillan), and Simon & Schuster.

¹¹ United States v. Apple, Inc. 13-3741-cv (L), United States Court of Appeals for the Second Circuit decision of June 30, 2015, at page 15-16.

¹² See Thompson (2012, chapter 1) for more detail.

¹³ See <http://authorearnings.com/report/november-2015-the-uk-report-author-earnings-on-amazon-co-uk/>.

identifies that while there are many similarities between the UK and US findings, the share of self-published books sold on Amazon via its UK website (amazon.co.uk) is significantly lower than that sold via its main website (amazon.com), while the share of the Big Five traditional publishers is significantly higher. There is one important factor which is likely to explain the difference between sales on Amazon's US and UK websites – the difference in pricing structures on the two sites. Big Five books are, on average, significantly more expensive in the US, while self-published indie books are cheaper in the US than in the UK. One reason for this is that since 2014, the Big Five have regained more control over final retail prices in the US. Following the end of the temporary ban on the agency model (i.e. publishers set the final consumer price, which arose from an antitrust action brought against Apple)¹⁴ publishers have negotiated a return to this model of pricing.

More recently a number of e-book subscription platform have emerged, essentially moving retailing from selling downloads of e-books to offering streaming services, which in many ways look like a library with an entry fee. In return for a monthly fee, readers can read as many books as they like from within the repertoire held by their chosen service.¹⁵ The two main subscription services are Kindle Unlimited¹⁶ and Scribd.¹⁷

It is important for the potential for self-publishing that, Amazon, in particular, had experience of dealing with small independent firms selling through its platform (using Amazon Marketplace) and, hence had a willingness to deal with and help smaller publishers including self-publishing authors. Therefore, if authors could somehow replace the inputs required to transform a manuscript into a book, which were traditionally provided by a publisher, they would have a route to the market and, therein, to the reading public. The arrival on the scene of new internet-based firms has enabled this.

2.3 *The entry of new services – disintermediation and digitalisation*

The wealth of information and new technology that has been made possible by the internet may give authors the means to bypass the traditional publishers and, potentially, side-step new and traditional retailers by dealing directly with those interested in reading their works. The

¹⁴ United States v. Apple, Inc. 13-3741-cv (L), United States Court of Appeals for the Second Circuit decision of June 30, 2015

¹⁵ Note that this is not a settled landscape. For example, in e-book subscription platforms, between July and September 2015, two of the four platforms, *Entitle* and *Oyster*, shut down.

¹⁶ Full information available at: <https://kdp.amazon.com/help?topicId=A156OS90J7RDN>

¹⁷ Authors need to go through Smashwords, INscribed Digital, BookBaby, or Draft2Digital to upload their work. Full information available at: <https://www.scribd.com/about>

incentive for this is clear when we consider information provided by the OECD (2012) about the likely distribution of revenues across the key actors in Figure 1 above. In particular, the share of the revenue going to the author, from which they have to pay any agent they may have, is calculated to be 8-15%. These are averages and we would expect this share to differ dramatically between authors. The net shares of revenue going to the publisher and the retailer are reported to be 30% and 40% respectively.

To by-pass the traditional publishers, authors may utilise one of the different platforms available such as Smashwords, Draft2Digital and Bookbaby all of whom offer various forms of assistance with the publication and distribution process. These services prepare the author's work for sale via the various retailers in that are forms compatible with the various readers. Authors also have the option of offering their book in paperback format with the help of print-on-demand services such as LuLu and Amazon's CreateSpace.

3. Legal rights in the UK and their effect on printed books¹⁸

The publishing industry and copyright law share a long and tumultuous history, but it is widely accepted that they also share an inter-linked, symbiotic relationship (see for example: Goldstein, 2003; Grosheide, 2001; and Wiseman, 2007). Copyright law aims to promote cultural diversity¹⁹ by providing creators with an incentive to invest in the creative process by offering legal protection which can be enforced to prevent unauthorised copying of works for a finite period. In this way, the author enjoys a lead-time to recoup their investment by having an option to exploit their work commercially. Despite these aims, UK copyright law has often been criticised for being more publisher-centred, than author-centred [see D'Agostino (2010, p. 53)].

In the past, publishers have had the important role of guaranteeing that the content is not libellous, obscene or blasphemous. It is clear that having a powerful gatekeeper who in return can be sued for breach of these rules is beneficial to policy makers and this in turn may go somewhat towards explaining why developments in the law have been skewed towards publishers.

¹⁸ This section focuses on the UK. European copyright laws differ, and collecting societies, standard contracts, and various forms of literary policies play different roles in different markets.

¹⁹ Through its paradigm, copyright law provides a legal remedy to increase the quantity and variety of cultural works produced and disseminated. This is likely to be increased with the reduction of transaction costs linked to self-publishing and the resulting higher competition in the book industry.

3.1 *Economic rights*

If a work satisfies the protection requirements, copyright law grants the author a bundle of exclusive rights.²⁰ These rights define the economic uses of the work, and hence the revenue streams, which the rightholder is able to control.²¹ The term of protection varies according to the type of work, but for a literary work, such as the text of a book, protection is fixed at seventy years after the death of the (last surviving) author.²² In summary, the CDPA grants the rightholder the exclusive right to reproduce the work by making any copies to distribute or otherwise communicate the work to the public (including placing the work online, as well as the exclusive right to translate or adapt the work [see Sterling (2015, pp. 442-445)]. The latter includes changing the format of a work, for example from a book to a film.

Generally, the law recognises the author as the ‘first owner’ of copyright.²³ So, while an author will be the first owner of their literary work, if a book is published by a traditional publisher, it is likely that the publisher will be responsible for the final layout of book. In this case, the publisher is the first copyright owner of the work comprising the additional protection for the typographical arrangement.²⁴ The question of ownership is vital in the book publishing industry given that, firstly, it is the starting point for the copyright term and, secondly, it continues to be essential for ensuring that national law respects the over-arching principle of national treatment as set out in international copyright law.²⁵

Enforcement of copyright is another aspect of legal protection which has been impacted significantly by the digital environment. Traditional publishers (to some extent rightly) feared that dissemination of works online would make it almost impossible to police their rights in the digital environment. Increased piracy (both one-off unauthorised uses and wholesale piracy) was predicted, as digital technology enabled works not only to be accessed, but for

²⁰ Section 16(1) CDPA and section 17-21 CDPA.

²¹ Alongside these economic rights, CDPA also grants additional rights to certain authorial works which qualify for copyright protection. These additional ‘moral rights’ are vested in the author alone, to protect certain non-economic interests of the author. Irrespective of whether the economic rights in a work are transferred, the author retains these moral rights which ensure that the author, nevertheless, retains some degree of control over subsequent uses of their works. Moral rights also further the general public interest, insofar as they protect the integrity of works by ensuring that the public is exposed to a work as intended by its author. Moral rights additionally reassure the public that any work is properly attributed.

²² Section 12(2) and 12(8) CDPA.

²³ Section 11 CDPA. Exceptions to this are made, for example, in the case of works created by employees.

²⁴ Sections 1(1)(c) and 8 CDPA. Sterling (2015, pp. 245 & 528).

²⁵ Sections 12 and 15A CDPA.

perfect, infinite copies of the work to be made and distributed at the click of a mouse.²⁶ However, the technological developments also created new opportunities for copyright holders. Seizing the opportunity to expand the exploitation of books to a new online market, publishers extended their offering to e-books, and devised ‘digital’ locks which deterred all but the most determined individuals from copying and printing the works made available online.²⁷ At the start of the digital publishing era, hardcopy books were scanned, encrypted and then made available to internet users in a downloadable format. However, once piracy in the form of bypassing encryption appeared, more sophisticated anti-circumvention measures were introduced and were reinforced by additional legal protection via copyright law, which rendered circumvention of rights management measures unlawful.²⁸ These new anti-circumvention provisions enabled right holders, for example publishers, to gain even more control, by making micro-management of every single use of a work possible (Wiseman, 2007). This shift signals the growing use of ‘mouse-click’ contracts and the like to regulate digital access to works, typically accompanied by the introduction of evermore sophisticated anti-piracy systems [Buydens and Dusollier, 2001 ; Galopin, 2012; Geiger, 2004; and Guibault, 2002]

3.2 The Importance of Contract Law

In publishing, the most common model is the ‘standard form’ contract model in which one side (here the publisher) dictates the terms on a take-it-or-leave-it basis to the author [Guibault

²⁶ There is a general debate as to whether reducing piracy has an effect on legitimate sales. For books from a field experiment carried out in Poland, Hardy et al. (2014, p. 17) “find no evidence of a significant shift in sales due to pirate copies being available on the internet”. By contrast, Reimers (2016) in a more extensive study does find a positive effect from piracy protection on legal sales of books but that the effect depends on various factors, including the format of the edition. “While physical formats are not affected by piracy protection, closer substitutes for online piracy such as legally distributed e-books see a mean differential protection-related increase in sales of at least 14 percent.” (Reimers, 2016, p 414).

²⁷ These digital locks (better known as digital rights management [DRM] locks), restore territorial frontiers in a way by enabling publishers and distributors to divide the book market in an online environment. DRM is used for: “(1) controlling access to and copying of reproductions and (2) administering information relating to the exercise of rights (e.g. in royalty collection and distribution systems).” Sterling (2015, p. 208). As Vaver notes, his system leads to an intricate situation whereby publishers lobby the legislatures arguing that the digital world has changed everything which requires more legal protection but simultaneously, if authors go to publishers to get a greater revenue share linked to the distribution of works in different formats, these claims are usually rejected based on the fact that the digital world has not changed the conditions (Vaver 2006, p. 6). This has also consequences on the tax regime applicable to e-books. The CJEU will have to decide whether Council Directive 2009/47/EC of 5 May 2009 amending Directive 2006/112/EC as regards reduced rates of value added tax (OJ 2009 L 116, p. 18) is invalid on the ground that it violates the “principle of fiscal neutrality to the extent to which it excludes the application of reduced tax rates to books published in digital format and other electronic publications”. As e-books are regarded as a provision of services, these cultural works subject to standard rate VAT while paperback book benefit from a reduced rate. With the Proposal for the new Copyright Directive, the EU legislator seem to favour a differentiation in VAT rates.

²⁸ Sections 296-296ZF CDPA implementing article 6 of the Infosoc Directive. For more on article 6, see Becker et al. 2003, p 463).

(2002, p.198)]. For example, many publishing contracts are wide-reaching and drafted by the publishers in their favour, covering the worldwide unlimited right to exploit the copyright-protected work in any form for the full copyright term. This position is exacerbated by the fact that many publishing houses adopt similar standard terms. Therefore, an author unwilling or unable to accept these standard terms was, at least traditionally, unable to get their work published however meritorious the work might be.²⁹

In this context, self-publishing is not only attractive for authors who wish to remain in control of their rights but the digital world has rendered successful self-publishing possible. Authors are now in a position to reproduce and distribute their works online directly to the public without having to bear printing, storage and transport costs. In addition, they can control the translations made of their works to preserve their integrity. Finally, in some instances the author is better placed, having regard to personal connections, to negotiate deals for the adaptation of the book into a film or a TV series.

In light of the above, self-publishing represents a feasible possibility for authors having a minimum level of technical competence who prefer to retain control of their intellectual property rights. The overview of copyright law demonstrated that the current copyright paradigm is still largely publisher-oriented. The over-arching principle of freedom of contract in UK contract law strengthens the position of publishers, since their superior negotiating position means they are often able to impose their terms on the author. The advent of new technologies (e.g. anti-circumvention measures) only served to strengthen the publishers' position, leaving little control and bargaining power for authors. While the law recognises 'moral rights', the manner in which these have been implemented within UK copyright law, such as the need for the attribution right to be specifically asserted to be effective, and the ability for other rights to be waived, undermines the practical impact on publishing practices.

4. The authors make, buy or delegate decision

The feasibility of self-publishing post-digitalisation implies that an author with a manuscript has to make a decision on whether to retain control or attempting to delegate the development of the manuscript to a publisher. Where the author chooses to retain control, further decisions

²⁹ This description should not be taken as saying that there is no mechanism for the courts in the UK to redress the unfairness of an unbalanced contractual relationship. Indeed, UK courts have held contracts voidable where there was a manifest inequality of bargaining powers, but there are several hurdles for the weaker party. *Peer International Corp v Termidor Music Publishers (No 1)* (2003) EWCA Civ 1156, (23); *Lloyds Bank Ltd v Bundy* (1974) 3 All. E.R. 757, (765); *Clifford Davis Management Ltd v WEA Records Ltd.* (1975) 1 All. E.R. 237, (240) (Denning LJ citing Diplock LJ in *Schroeder Music Publishing Co Ltd v Macaulay*); *Schroeder Music Publishing Co Ltd v Macaulay* (1974) 3 All E.R. 616, (624); *Fry v Lane* (1888) 40 Ch D 312 (Kay J).

are necessary because there are many aspects of the publishing process, such as editing, file translations and cover design, which the author could either do or delegate to one of the new intermediaries. While there are many anecdotes about how authors have made these decisions, there is little more general evidence. This is an area where good accessible data is in short supply. In this section, we rely on a data set from a non-representative survey of authors and the reports of results from another survey, where the representativeness is hard to gauge. This clearly implies that we need to take the results found in this section with a large pinch of salt. However, common to both samples is a likelihood that active self-publishers are over-represented. Observing their choices can still tell us something about the behaviour of modern authors.

4.1 The data

To gain some insight into the choice made by authors, we use a dataset arising from an online survey carried out by *Author Earnings*.³⁰ At the time we downloaded the data, a total of 1,704 authors had responded to the survey.³¹ In responding to the question: "Based on primary earnings, what kind of author are you?", 264 declared that they were "traditionally published" and 1440 declared that they were "self-published".³² While this indicates that the sample is not representative of the pool of published authors and that in particular self-publishing authors are very overrepresented,³³ nonetheless we believe that we can learn something about what may motivate the choices of authors.

4.2 The "do or delegate" decision

In the sample, the respondents who classify themselves as self-published do not have as much long-term experience of the industry.³⁴ While roughly a third of the traditionally published authors published their first book before the advent of the Kindle in 2007, the corresponding

³⁰ Dataset and report are available at the following link: <http://authorearnings.com/results/>. The website was visited on 18 February 2016. The data starts in 2014, several years after self-publishing became a credible choice in the market. This is an obstacle to the analysis of publishing trends, but it allows for a detailed snapshot of the market.

³¹ It is important to note that this is an open survey to which people continue to add responses. For this reason alone, this is not a representative survey. Furthermore, since the first respondent took the survey on the 8 February 2014, while the last in our sample took the survey on 18 February 2016, we cannot guarantee that the same person did not take the survey several times. We have also removed a number of outliers where authors provided extreme responses, such as estimated total earnings from traditionally published books of \$160M or publishing in excess of 200 books.

³² Of the latter, 20 also declared that their primary publisher, based on earnings, was either one of the Big Five or a smaller publisher.

³³ Data on the number of authors who self-publish and the number of books self-published is very poor.

³⁴ Those authors with less than one year of experience have been eliminated.

figure for self-published authors was only about 7%. This confirms a skew towards new entrants into book writing, and this must be kept in mind when interpreting later findings. Of the authors in the sample who have a long track record, there is a disproportionate number with experience of traditional publishing, because they would not have had the option to self-publish when they published their first book.

One of the questions asked in the survey is: “How do you plan on or hope to publish your next work?”³⁵³⁶ The answer allows the sample to be split into four subsamples. The largest subsample, 1,366 respondents, consists of authors who declared that they are currently self-published and plan on remaining so. The second largest, 159 respondents, consists of authors who declared that they are currently traditionally published and who plan on remaining so. However, there are also subsamples where the author wishes to switch from traditional to self, 88, and from self to traditional, 66. Of the 1,421 who were content with their current mode of publishing, 338 or 22% had experience of both forms in the past. Interestingly, a larger fraction of the 154 who would like to switch in the future, 44%, have prior experience of switching. This increased fraction of experienced authors among those wishing to switch is driven entirely by the fact that 60% of those who are published traditionally have self-published in the past. Hence the majority of authors who are planning a switch to self-publishing are taking an informed decision based on past experience, while the majority of those planning a switch from self-publishing to a traditional publisher are basing their decision on expectations.³⁷

Decisions about which mode of publishing to choose in the future depends on both options and experience. While we cannot infer anything about options, the data does offer a glimpse of the experience of the different groups. One set of questions relates to the “performance” of the respondents. The averages for the four groups are reported in Table 3 below. To focus on switching vs staying, we have blacked out the cells asking traditional published (self-published) about their experience related to books they may have self-published (traditionally published).³⁸

³⁵ Note that there is a degree of ambiguity in the question. While “hope to” implies a choice, “plan on” could be by necessity.

³⁶ This question is answered by 1679 of the respondents, leading to a loss of eight self-publishers and 17 traditional publishers.

³⁷ In the working paper version (<https://zenodo.org/record/321609/files/CREATe-Working-Paper-2017-06.pdf>) we report out more formal econometric analysis which confirms the correlation between success in terms of being able to earn a full time living off their writing and having experience of alternative modes of publishing. This may simply indicate that more entrepreneurial authors tend to do better.

³⁸ The values in those cells are positive but small.

[Table 3 around here]

Note that in all cases, those who wish to switch are on average doing less well than those who plan to stay with their current method of publishing.

The desire to switch may be caused by both a pull and a push factor.

A second set of questions relates to what aspect of either traditional publishing or self-publishing they are satisfied with and offers some insights into this. The options for answers reflect what is often assumed to be the main benefits or challenges of a particular form of publishing and are presented in Tables 4 and 5 below. For traditional publishing, the issues are – broadly speaking – centred around the quality of editing, the income generated, the effectiveness with which the books are distributed and the working relationship with the publisher. For self-publishing, the issues are about the new additional tasks the author takes on: the quality of the final product, how much effort has to be put into sales, how well the book is distributed and how many copies are eventually sold.

[Table 4 around here]

Those who prefer to continue to publish traditionally are satisfied with the services they receive and much more so than any of the other subsamples. This is particularly the case with editing, something which has been found in another recent survey.³⁹ The exceptions relate to income generation. They may be (just) more satisfied than not on average about the price at which their book is sold, but not with the promotional effort and the resulting sales.

As one might have expected, those who wish to switch to self-publishing are less satisfied across the board and, more interestingly, only the editing support received more than 50% approval. This suggests more of a push factor and, in particular, that the authors who wish to switch are very unhappy with the income generated (they are unhappy with sales, price and promotions) and with the way the publisher handles e-book sales.⁴⁰ Some of those who see themselves as self-published have experience of traditional publishing. While it is easy to dismiss the negative attitude of those who want to continue to be self-published as arising from self-confirmation bias, note that even the small number with this experience who want to switch are negative about the support they received when they published traditionally, though not

³⁹ A more recent survey commissioned by authors Jane Friedman and Harry Bingham, mainly covering traditionally published authors, finds that more than ¾ of respondents found both the editorial input and the copy editing of their publisher to be good or excellent, see Q7 and Q8 in report available at <https://janefriedman.com/author-survey-results/>.

⁴⁰ A weakness of the data is that we do not know if the author had a real choice to remain with a publisher. If not, the reported unhappiness may simply reflect an author justifying a forced change – the push is imposed.

nearly as many as those who are currently self-published and who wish to remain so. However, despite this negative experience of support, they are still ready to switch.

Turning to the level of satisfaction by those who have experience of self-publishing, in Table 5 below, across the board, every group is on average very satisfied with the overall quality of the final book.

[Table 5 around here]

Concerns about how “good” the final product looks does not seem to be important in determining the mode of publishing. Among those who self-report as self-published, switchers are otherwise less satisfied and presumably hope life will be better after the switch. For those who self-report as traditionally published, switchers have a more positive memory of self-publishing, and this may be what motivates the switch.

4.3 The “make or buy” decision

When choosing to self-publish, a number of tasks, which traditionally would have been carried out by the publisher, such as editing and cover design, fall initially on the author. However, as illustrated in section 2.3, some or all of these tasks can be delegated to others. The questionnaire asks two questions, which relate to how self-published authors deal with the important tasks of editing and the designing the cover art. The responses are covered in Table 6.

[Table 6 around here]

Note first of all that very few authors in either subgroup rely solely on themselves to do the editing with no added help. Secondly, among those who plan to self-publish next time, many more expect to use professional services to deliver editing. In particular those who have past experience of self-publishing and who want to switch to that mode are seemingly prepared to hire professional help. Among this particular sample, self-publishers do not necessary want to retain control over all aspects of publishing. However, cover art appears to be an area where those who plan to stick with self-publishing are more self-reliant than either group of switchers. If we focus on those who hired professional help both with editing and cover art, overall 36.5% of the sample did so. Three of the four subsamples are very similar while the non-switching traditionally published author subsample has significantly fewer authors who use professional help with self-publishing. ⁴¹

⁴¹ In the working paper version (<https://zenodo.org/record/321609/files/CREATe-Working-Paper-2017-06.pdf>) we report out more formal econometric analysis which a significant positive correlation between being a successful self-published writer and hiring an editor – no such correlation was found between success and

5. Publishers' power to select and promote

Designing the best marketing strategy for a book is complicated by three factors. One is that a book can be categorised as an 'experience good' (Nelson, 1970). A second is that since consumers' tastes differ, one consumer will not necessarily agree with an earlier reader's rating of book. This makes the recommendations and reviews of others an imprecise guide, especially if nothing is known about these reviewers. A third is that previous books by the same author are only an imperfect guide to how much the reader will enjoy other publications by the same author. Thus, when choosing a book to purchase, a consumer will base this decision on imperfect information drawn from their own past purchases, from endorsements and professional reviews and from the comments of others who have recently purchased that book.

The first and third factors have not been changed dramatically by the increased digitalisation, but the second has. With the increased importance of reader reviews, an author wishing to maximise sales of their work faces a fundamental problem not only how to get their book noticed and talked about, but also to ensure that it comes to the attention of the 'right' people: its target audience, who will value the work most. Several studies have demonstrated that word of mouth recommendation and positive reviews are a key determinant to the success of various experience goods (Anderson and Madruger, 2012; Beck, 2012; Chevalier et al., 2006; Lee et al., 2015). The internet greatly enhances this spread of product information (Baye et al., 2015). Consumers can use this tool not just to track down opinions and reviews written by professional book reviewers, but also those provided by fellow consumers to obtain advice based on actual purchasing decisions.

While publishers, armed with their greater experience and access to funds, may be able to promote a book effectively,⁴² they still need to undertake cost-benefit analysis to identify the type and amount of publicity any given manuscript is likely to need to be a commercial success, and to evaluate whether sales are likely to be sufficient to warrant this upfront expense. This is far from straightforward. In traditional publishing, manuscripts which proved to be successful have often already been rejected by other publishers in the past, especially where the manuscript is by a first-time author. A new author with a superior product, but without a proven

purchasing cover art. We also found a negative correlation between the probability of switching from self to traditional publishing and hiring an editor. If not having an editor makes it less likely that the self-publishing author is successful, then this could explain some of the motivation for the desire to switch.

⁴² While intuitively one might expect publishers to be always able to outperform authors, this is not generally true. Carolan and Evain (2013) provide some illustrative case studies where marketing is driven by the author.

track record, risks being overlooked by the publishing industry⁴³. Secondly, some books have been missed altogether and first gained prominence through the author self-publishing their work. The success of the *Fifty Shades* trilogy demonstrates how a self-published book which has received online acclaim can then be translated into a bestselling print book.⁴⁴ It is also a testament to the broad consensus among self-published authors that “traditional publishers can accomplish what they can’t, including distribution, marketing, and selling foreign translation rights”.⁴⁵ In light of these two challenges, this section questions what effects the increased ease of self-publishing might have on publisher behaviour at the manuscript evaluation stage.

Imagine a situation where there are several competing manuscripts submitted to the publishers, each within the same genre. If published, any of these manuscripts on their own would command reasonable sales, but if all were published, one would stand out in terms of sales. In a pre-digital world, books can only be published successfully through an established publisher. Such a publisher might have little incentive to expend effort in identifying the most commercially viable book of a sub-genre unless it feared that this manuscript might end up with a rival publisher. If the total number of manuscripts on offer is large, the likelihood of a publisher facing competition in this one particular category before most of the demand had been met is relatively small.

In a pre-digital world, the cost for publishers of “getting it wrong” may not have been significant. In a digital world, any author can self-publish using a process, which is reasonably quick and inexpensive. The internet increasingly creates mechanisms, which enable enough of the potential readership to assess the relative quality of any book, irrespective of whether it is traditionally published or self-published. In this world, a publisher’s mistake in selecting the less commercial book for publication has potentially more severe consequences: the author of the better, but rejected book, has the potential to benefit significantly from self-publishing, and

⁴³ A dramatic example of this is the *Harry Potter* series, one of the most successful series of books in the industry, and particularly in the UK. Even Bloomsbury, though willing to take a chance and publish the book, were surprised at its ultimate success, because they failed to identify the manuscript as any more promising than any of their other children’s books. The case is testimony to the difficulty even professionals have in assessing the commercial merit of a manuscript.

⁴⁴ Some other successful self-published books are: Hugh Howey’s *Wool* (2012), which sold over half a million copies through Kindle Direct Publishing; Andy Weir’s *The Martian* (2011) which was turned into a successful film, and *The Kissing Booth* (2012) written by the 17-year-old Beth Reekles who later signed a million-dollar three-book deal with Random House’s Delacorte imprint.

⁴⁵ Sales, B. (2013). “Fifty Shades of Grey: the New Publishing Paradigm”, *The Huffington Post*, 18/04/2013. URL: http://www.huffingtonpost.com/bethany-sales/fifty-shades-of-grey-publishing_b_3109547.html.

such a decision has an adverse effect on the rejecting publisher through increased competition.⁴⁶

This leaves the publisher with two options. The first is to devote more effort to scrutinising new manuscripts to increase the probability that the best book is correctly identified. Alternatively, the publisher can take a step back, accept that some authors will self-publish, use the market for self-published works to identify the ‘best book’, and then approach the author with a view to taking the publication onto the next stage. This section contains a stylised illustrative model that identify some of the effects on publishers arising from the threat for self-publishing.

5.1 The formal model

Consider two books, X and Y, by different authors but in the same genre, with significant latent demand. We will assume that one of the books strictly dominates the other in the sense that if both were marketed at the same time, one would attract a disproportionate amount of the aggregate demand. This will be referred to as the “better book” in the following, where ‘better’ refers to its ability to generate revenue. To keep matters simple, we assume that each book manuscript is a priori equally likely to lead to the better book. We want to allow for the book to be published through a traditional publisher or self-published. For simplicity, we will assume that there is only one publisher.⁴⁷ Demand is summarised as follows where, for simplicity, we treat price, P , as exogenous:

	Probability	Demand if only one is published	Demand if both are published	
			Demand for X	Demand for Y
X is the better book	1/2	D_X	$D_{X Y}$	$d_{Y X}$
Y is the better book	1/2	D_Y	$d_{X Y}$	$D_{Y X}$

We assume that $D_{i|j} + d_{j|i} > D_i > D_{ij}$ where $i, j = X, Y, i \neq j$, which amounts to assuming that even if i is the better book, if j is also available, some sales of i are lost to j ($D_i > D_{ij}$) and some people want j and not i so that demand is expanded by j 's presence ($D_{i|j} + d_{j|i} > D_i$). Further, we assume that while the loss in demand from choosing the less good book is small if

⁴⁶ If publishers respond to this by offering smaller advance payments, the incentive for each author to seek out a publisher is much diminished. Consequently, the best books might never be offered to publishers first.

⁴⁷ While this is extreme, given the sheer volume of manuscripts, the chance that the two books are discovered by two different publishers at the same time is small. For example if we consider the case of the UK, according to the 2013 International Publishers Association (IPA) report, in 2014 the UK published 184,000 new and revised titles in 2013, that is 2,875 titles per million inhabitants.

only one book is published, i.e. $|D_X - D_Y|$ is small, the loss in demand from publishing the less good book if both books are published is very significant, i.e. $|D_{i|j} - d_{i|j}|$ is large. Finally, we assume that $D_{i|j} + d_{j|i} = D_{j|i} + d_{i|j}$, i.e. total demand for the books when both are published is independent on which is the best. The second set of assumptions ensures that the potential loss from making a mistake is significant.

We will assume that the cost for a publisher to develop a book is F , while the cost for a self-publishing author is f . These costs include producing and marketing the book. One would expect there to be a difference between F and f , with $F \geq f$ as the optimal marketing campaign for the publisher to be more effective and hence, in equilibrium, would be allocated a bigger budget. In line with that, we assume that the author can only realise a fraction, μ , of the total potential demand. Finally, the publisher has to pay an advance, A , to any author it signs up.⁴⁸ For simplicity, we assume that the publishers have all the bargaining power so that A equals the authors expected value of their next best alternative.

Consider the following order of events:

1. The authors simultaneously complete a manuscript of unknown commercial value
2. The publisher review the manuscripts and picks one to read at random.⁴⁹ Reading one manuscript does not reveal any information other than an expected commercial value. Without reading both scripts, they are expected to have the same quality
3. The publisher either
 - a. Makes an offer to the first manuscript, w.l.o.g. manuscript X.
 - b. Identifies and reads the second script. This fully reveals all values in the table to the publisher. An offer is made to the best book.
 - c. Rejects X and does not search for Y
4. The author observes whether their manuscript has been rejected and whether a competing book has been published by a publisher.⁵⁰ The authors cannot observe the action of the publisher, i.e. which manuscript(s) have been reviewed. If the publication

⁴⁸ If the publishing house decides to publish the book, it acquires the manuscript from the author and negotiates the future royalty rate to be paid to the author including any advance on future royalties

⁴⁹ We explicitly rule out the strategy where the publisher simply signs up both manuscripts. If μ is very small and/or f large, the expected profit from self-publishing is small and it would be relatively cheap to acquire both manuscripts and if F is also small, this may even be worthwhile. We offer two motivations. Firstly, we are interested in cases where self-publishing is a realistic strategy. Secondly, any result that it is worthwhile to essentially foreclose the self-publishing market would disappear once the number of publishers and/or manuscripts were increased.

⁵⁰ A possible justification is that such publications are typically announced well in advance of the publishing date.

of a rival book is announced, the author of the other manuscript can assess the relative merit of the two books.⁵¹ Any rejected or undetected author decides whether or not to self-publish.

5. If no manuscript has been accepted by the publisher, it can wait until self-publishing reveals which book is the better one and then make an offer to the author. Because it takes sales and reader reviews for the quality to be revealed, remaining demand is reduced by S , the amount self-published copies already sold.⁵²

An implication of step 4 is that only if the other manuscript is published will an author know the true level of demand, i.e. the values in the table above. If not, decisions have to be based on expected values. Let $E(\cdot)$ be the expectations operator. Given the symmetry of the model, we assume that $E(D_{X|Y}) = E(D_{Y|X})$ and $E(d_{X|Y}) = E(d_{Y|X})$.

To rule out cases where self-publication always dominates not publishing, we assume that if someone is committed to publishing the better book, self-publishing the other book is never profitable,

$$P\mu d_{j|i} - f < 0 \quad (1)$$

However if the less good book is being published, self-publishing the better book is profitable:

$$P\mu D_{j|i} - f > 0 \quad (2)$$

The key implication of assumptions (1) and (2) is that if the publisher fails to identify and publish the better book, the author will choose to self-publish. The consequences for the publisher are that if he searches for the better book, there will be no competition whereas if he does not, and manuscript X is not the better book, he will face competition.

To calculate the minimum fee for a book manuscript, we will make one further assumption. If no book is published by a publisher, the expected return to an author from self-publishing even if both books are published, is positive. Formally:

$$\Pi_i^a = \frac{1}{2}(P\mu E(D_{i|j}) - f) + \frac{1}{2}(P\mu E(d_{i|j}) - f) > 0, i, j = X, Y, i \neq j \quad (3)$$

⁵¹ The somewhat artificial information revelation simplifies the analysis considerably as no signalling is implied from the action of the publisher. An alternative assumption is that self-publishing is quick so that having actually read the competing book, an unpublished author can evaluate the book and publish their own book without losing much future demand.

⁵² A justification for this is that self-published books are only reviewed by readers and hence it takes a minimum number of sales to accumulate enough reviews that the quality (i.e. latent demand) is revealed.

Note that given our symmetry assumptions, $\Pi_i^a = \Pi_j^a$. An implication of (3) is that, if an author is offered a book deal and if self-publishing is possible, the alternative is to self-publish with the profit given by (3). In this case, an author has to be offered at least a fee $E(A_i) = \Pi_i^a$ for their manuscript to ensure acceptance.⁵³

Lemma 1: A publisher is more likely to search for the better manuscript when self-publishing is a credible threat.

Proof: If self-publishing is not a credible option, searching is more profitable if the expected increase in sales when Y is the better book outweighs the cost of search, i.e.

$$\frac{1}{2}P(D_Y - D_X) > C \quad (4)$$

If self-publishing is a credible option and the publisher does not search, expected profit is given by

$$\Pi_{NS}^p = \frac{1}{2}(PD_X - A_X - F) + \frac{1}{2}(Pd_{X|Y} - A_X - F) \quad (5)$$

While if the publisher does search, the expected profits are given by:

$$\Pi_S^p = \frac{1}{2}(PD_X - A_X - F) + \frac{1}{2}(PD_Y - A_Y - F) - C \quad (6)$$

Comparing (5) and (6), and recalling that $A_X = A_Y$, searching is more profitable if the expected value of gained net sales, as a result of searching, exceeds the cost of searching:

$$\frac{1}{2}P(D_Y - d_{X|Y}) > C \quad (7)$$

Comparing (4) with (7) it is clear that (7) is much more likely to hold which prove the lemma.

This result does not necessarily imply that publishers will become more active. The opposite could happen because of the assumption that if the publisher publishes neither manuscript, both authors will self-publish. This opens up an additional strategy for the publisher, a wait-and-see strategy, in which the publisher lets the market identify the better book and then uses the fact that it has greater ability to commercialise the book by signing up the better book once it has proven itself.

Assume for the moment, that if no manuscript has been accepted by the publisher, the authors decide whether or not to self-publish without knowing the decision of any possible rival. This simultaneity decision precluded any wait-and-see strategy on behalf of the authors. The

⁵³ The outside option when self-publishing is not an option has been normalised to zero. As Peukert and Reimers (2018) point out, a consequence of the outside option of self-publishing is that the outside option when negotiating with a publisher has gone up, increasing the expected fee, and one would expect, the overall supply of manuscripts.

consequence of sequential decision-making will discuss below. Equation (3) ensures that both authors would self-publish if no publisher has picked up the book. If an author is offered a book contract after self-publishing and selling S copies of their book, they will know that theirs is the better book and that their future sales are $\mu D_{ij} - S$. As the cost of publishing, f , is sunk cost, the fee to author i would have to be at least

$$A_i^{Was} = P\mu D_{ij} - PS \quad (8)$$

The fee the publisher expects to pay before looking at any manuscripts is then given by

$$E(A_i^{Was}) = P\mu E(D_{ij}) - PS$$

The aim is to understand whether such a wait-and-see strategy could ever be optimal for the publisher. To realise the full demand, we assume that the publisher has still to incur publishing and marketing costs, F . The next lemma demonstrate that this is possible

Lemma 2: If the cost of searching for and evaluating manuscripts is relatively high and the self-publisher relatively poor at realising total potential demand, the publisher will choose to let authors self-publish first and then purchase the rights to the more successful books.

Proof: From Profit to the Publisher if choosing search is, from (6) given by

$$\Pi_p^S = P\left(\frac{1}{2}D_X + \frac{1}{2}D_Y\right) - E(A_i) - F - C \quad (9)$$

While the profit from choosing wait-and-see is:⁵⁴

$$\Pi_p^{Was} = P\left(\frac{1}{2}D_{X|Y} + \frac{1}{2}D_{Y|X}\right) - E(A_i^{Was}) - F - PS \quad (10)$$

Note that Wait-and-See dominates Search if:

$$C - f > P\left(\frac{1}{2}(D_X - D_{X|Y}) + \frac{1}{2}(D_Y - D_{Y|X})\right) + P\mu\frac{1}{2}\left(E(D_{ij}) - E(d_{ij})\right) > 0 \quad (11)$$

The right-hand-side of (11) is positive, so a necessary but not sufficient condition for the inequality to be satisfied is that $C > f$. The first term on the right-hand-side is relatively small as we have assumed that the weaker book does not impose much of a competitive constraint on the better book. The second term is increasing in μ and zero if $\mu = 0$. Let $\Delta = P\left(\frac{1}{2}(D_X - D_{X|Y}) + \frac{1}{2}(D_Y - D_{Y|X})\right)$. Then for any $C > f + \Delta$, there exists $\mu^* > 0$ such that for all $\mu > \mu^*$, the equality in (13) is satisfied. The key parameters are then C and μ . In other words,

⁵⁴ In calculating this, we make two assumptions, both of which makes the wait-and-see strategy less attractive. One is that despite the author having expended f , the publisher still have to expend F to reach the bigger audience. The second is that we have not included any costs for finding and reading manuscript X in the first place.

for the Wait-and-See strategy to dominate, the cost of searching for and evaluating manuscripts has to be relatively high and the self-publishing author relatively poor at exploiting total demand. This proves the lemma.

Contrasting lemma 1 and 2, the model demonstrate how self-publishing could in theory either lead to more effort by the publisher or more initial self-publishing. Each strategy has a different impact on the best-seller lists. The first strategy should see a drop off in the number of self-published books entering the list; the second should lead to an increase in the number of successful titles which were originally or previously self-published.

An author is better off if more income is generated by the Wait-and-See strategy. For the author of the better book, this requires that $E(A_i^{Was}) + PS - f > E(A_i)$, which is true as

$$P\mu E(D_{ij}) > P\mu \left(\frac{1}{2}E(D_{ij}) + \frac{1}{2}E(d_{ij}) \right)$$

Thus while it is ambiguous whether the fee is larger when the publishers chose the Wait-and-See strategy, when combined with the value of initial sales, the author of the better book is better off when this strategy is employed. From an ex ante perspective, the expected income from the Wait-and-See strategy is $\frac{1}{2}(E(A_i^{Was}) + PS) + \frac{1}{2}E(d_{ij}) - f$ which equals $E(A_i)$. From an ex ante perspective, the author who does not know about the quality of their manuscript is indifferent between the two strategies. Finally, as the Wait-and-See strategy gives rise to more diversity, one could argue that the consumers are also better off with Wait-and-See strategy.

The model is clearly very stylised and has consciously avoided any signalling effects or information transmission arising from the behaviour of the publisher. More generally, the information revelation mechanism has been kept very simple. For an attempt to model this, see Peukert and Reimers (2018). Their model focuses on learning from observed sales and essentially formalises the process through which the Wait-and-See strategy leads to learning. We assumed that if no manuscript was accepted by the publisher, the authors simultaneously decided whether or not to self-publish. Whether there is a value to an author from a wait-and-see strategy depends on the cost of developing a loss making book which is avoided from compared to any lost sales from being second. To model this would require a more dynamic model, considering the effect on competing books on sales more extensively. One possible argument in favour of the simultaneous approach is that it is easier to carry out the evaluation of relative merits when the alternative is a professionally published and marketed book.

5.2 *Testable implications*

Each strategy has a different impact on the best-seller lists. Increased search and closer scrutiny of manuscripts should ensure that the traditional publishers maintain their role as the gatekeepers of the publishing business and give rise to only few self-published books entering the best-seller lists. Using self-publishing as an information revelation mechanism should lead to an increase in the number of successful self-published titles, at least temporarily until the author is signed up, but possibly permanently as some authors may get a taste for remaining in control. The strategies also have an impact on the long tail, which is separate from the straightforward effect of an increased number of books in the tail arising from the much lower cost of self-publishing. Compared to the “do nothing” strategy, either means of learning about the quality of a manuscript should lead to the major publishers having fewer publications in the tail, i.e. fewer flops. More generally, it should lead the major publishers to publish fewer books, leaving more manuscripts to be published by smaller publishing companies as well as by the authors. Importantly, this should manifest itself in major publishers experience a larger drop in share of unit sales than revenue. The Wait-and-See strategy, whether employed by the publishers or the authors, should encourage more self-publishing and in particular more investment by the author in promoting their book, because the expected benefits in this case is higher irrespectively of whether they continue to self-publish or sign an agreement with a publisher.

The availability of data is a major obstacle in studies of the book industry (Waldfogel and Reimers, 2015) and we do not have ideal and readily available data to test any of the implications of the theory noted above. Such data would require knowing the history of book manuscripts from idea to entering a bestseller list and having detailed data on the division of earnings from books. In place of this, we have some data sources which are based on a sampling methodology which has not been tested, some data sets which are based on samples which are either evidently biased or possibly biased and a large number of anecdotes and casual empirical observations culled from the internet. Some of these sources relate to a particular jurisdiction, often the US, but some do not. Combining these sources may still enable us to tell a story which may or may not support some of the proposed strategies.

An increase in sheer numbers of self-published books is in itself not enough to indicate that the wait-and-See strategy has gained traction. Some of the past *Authors Earnings* reports have been looking at how well represented self-published books are in best-seller lists and in the best-earnings lists. Looking at data from the US across seven quarters, from February 2014 to

September 2015, *Author Earnings*⁵⁵ considers the distribution of authors who are consistently⁵⁶ earning at least \$X per year across four modes of publishing: (i) self-publishing (referred to as indie), (ii) Amazon published, (iii) traditionally published, and (iv) hybrid,⁵⁷ and conditioned on when the author published their first book. Interestingly, the pattern remains the same whether \$X is \$10K, \$25K, \$50K, \$100K or indeed \$1M, which suggests that the insights are fairly robust. The only two categories, which really matter, are self-published and traditionally published. For X = \$10K we summarise the share of self-published authors in Figure 2 below.

Figure 2 about here

The share of self-published authors increase steadily as we move from those authors who published their first book sometime in the past century through to those who started publishing after 2013. While those who self-publish account for less than 40% of all authors with that level of income, when restricting the sample to the most recent group recorded, this fraction is over 60%. The consistent picture emerging from their data is that, among those who have published their first book within the last five years, the majority of those who have been successful in selling e-books were self-published.⁵⁸ This is interesting as these are authors who had a meaningful choice between traditional and self-publishing from the beginning of their careers. Obviously, it would be good to have these insights verified by other sources of data, but the ability of, at least some, self-published authors to perform well suggests that publishers are not the only gatekeepers.

Authors Earnings January 2018 report looks at a more comprehensive set of data than previous reports and covers the three quarters Q2 – Q4 2017.⁵⁹ From our perspective, it is interesting to note two observations. Firstly, self-published indie authors feature significantly on their list of top selling author by \$, accounting for 7% of the top 100, 20% of the top 250, 24% of the top 500 and 28% of the top 1000. Secondly, they found that there had been a large change in the

⁵⁵ Report: “Individual author earnings tracked across 7 quarters, Feb. 2014 – Sept. 2015”, available at: <http://authorearnings.com/report/individual-author-earnings-tracked-across-7-quarters-feb-2014-sept-2015/>. It is important to note that these are authors earning more than \$XK per year consistently from only that subset of their Kindle books that appear on the Amazon best seller lists. Thus, this clearly underestimates authors’ total earnings.

⁵⁶ By which they mean that the author has earned at least \$X for two of the seven quarters. Their hope is that they have excluded the “one-day wonder”.

⁵⁷ An author is classified as ‘hybrid’ if at least 25% of the author’s earnings come from more than one publishing route.

⁵⁸ An alternative explanation is that this is simply a result of more books being self-published. If the number of successful books increase with the number of books written, then mechanically, if the fraction of books published by self-publishing authors increase over time, then so would the fraction of successful self-published books.

⁵⁹ <http://authorearnings.com/report/january-2018-report-us-online-book-sales-q2-q4-2017/>

list of top earning indie authors, so that a lot of the top-selling indie authors were relatively new names, while a lot of the pioneering indie superstars did not make the new list. It appears that self-publishing authors are becoming more attuned to the new medium and that it may be them, who are choosing a wait-and-see strategy.

There may be other reasons for publishers to move quickly to sign up new talent and for authors to self-publish. Film and TV producers also appear to be monitoring self-publishing authors. According to an article in the Guardian,⁶⁰ “Bestselling self-published authors attract producers because they have a proven track record if they stay on Amazon sales charts over time”. In the same article the importance of contracts and contract terms is also spelled out, see section 3.2 above.

Even if one has some reservations about the quality or reliability of the data from Author Earnings, the threat posed by self-publishing to the gatekeeping role and centrality of the traditional publishers appears credible. While the wait-and-see and the self-publishing strategies place more risk on the individual author and forces them to become more entrepreneurial, the reduction in the costs and complexities of self-publishing and the support available from new intermediaries and retailers suggests that authors can and do take more control. The availability of a credible outside option is likely to benefit authors more generally, including those who have chosen to be traditionally published.

6. Future problems: the long tail and new gatekeepers

Above, we have argued that authors are likely to benefit financially from digitalisation, either directly through self-publishing or indirectly through negotiating better terms. However, this may be tempered by a significant adverse effect from over-entry into the industry. This phenomenon, often referred to as the long tail,⁶¹ refers to the huge number of books that hardly sell any copies at all and are rarely reviewed either by professional reviewer or readers.

The large number of available books creates a problem both for the author and for the potential reader. In the past, publishers served as gate-keepers together with professional established book reviewers. The large increase in self-published books has not only undermined the role of the publishers but the position of reviewers who could never hope to cover even a fraction

⁶⁰ Danuta Kean, “Show me the money!': the self-published authors being snapped up by Hollywood”, The Guardian, Mon 15 May 2017, available at <https://www.theguardian.com/books/2017/may/15/self-published-authors-hollywood-andy-weir-the-martian-el-james>.

⁶¹ See e.g. Brynjolfsson et al. (2006, 2010).

of books published. While some have welcomed a world without gate-keepers, they have also pointed out that such a world is not for all.

The need for “gate-keepers”, or at least information aggregators, arises because books are experience goods where sampling is costly in time because the quality of the book is not known for sure until the very last page has been read and, even then, potentially not very informative (the author may be a one-hit wonder). Sampling a chapter may help a reader to decide whether they like the writing style, but is not determinative of whether the author is able to deliver an appropriate plot or captivate a reader for the entire length of the book. Moreover, for most books and most readers, the book will only be read once. In that way, books differ from other products provided by the creative industry, such as music and video games, where much can be learned by sampling and where the product is likely to be used more than once. One consequence is that, as consumers, we experiment very little, selecting works from tried and tested authors, or adopting the view of others, relying either on traditional reviews by “experts”, by the mass of customer reviews enabled by the internet, or by the recommendations from trained book store staff [see Chavelier et al. (2006)].

Searching for books is simultaneously facilitated by the internet and disabled by the “needle in a haystack” phenomenon, particularly as it is easy to lose confidence that you will recognise then needle when you see it. In all of this, the author and the reader face a linked problem: how can an author make himself visible so that their work may be found? This problem may be solved for some consumers by evermore sophisticated internet services, but it is unlikely to be solved for all. This presents a problem with achieving the policy objective of cultural diversity which is likely to exacerbate existing problems with access for all [see Shaver (2014)]. Strategies authors might employ to attract attention, such as setting very low or even zero prices, may well backfire because readers search less carefully when the book is (almost) free and hence the match between author and book is, on average, less good. A poor match may result in a negative review for that reason alone, an effect identified in Zegners (2016a and 2016b). Interesting, two new developments in the brick-and-mortar end of retailing may suggest that solving the search problem only through on-line platforms may be difficult. Firstly, Amazon has by now 15 bookstores, including the one that just opened in DC, and soon they will open one in a suburb of Idaho's state capitol.⁶² More generally, the number of independent

⁶² See <http://www.idahostatesman.com/news/business/article206393779.html>.

bookstores in the US, following a period of decline is now on the increase.⁶³ Whether this extends to the UK is at present unclear.

A problem with gate-keepers is that most readers would prefer only to consult one, and most authors would prefer to deal with only one. This preference for monopoly from the service perspective may store up problems for the future, especially if it is inevitable that for gate-keepers to be useful they also have to be powerful. In that case, we may just have swapped one group of powerful intermediaries with another.

7. Conclusions

This paper has shown how the book industry has been changed significantly by digitalisation once there was significant entry of e-readers into the market. The internet and in particular the development of the e-book has created new types of service providers and retailers. These challenge traditional publishers by enabling authors to completely by-pass them or at least threaten to do so. The likely effect of reducing the cost of entry for an author who can navigate the self-publishing maze is an increase in the number of works published. Arguably, the main question is, no longer, how can we ensure that the financial rewards to authors are such that an adequate number and variety of books are produced and published. Indeed, the incentives now appear too powerful and, just as disruptive technology in other markets has challenged how these markets have been governed legally, we may need to do the same for creative industries and the design of copyright laws.

With self-publishing, authors are able, at least, to remain the owners of their rights, meaning that authors are better-placed to exploit their rights themselves, or hire experts for specific roles. This does not mean that traditional publishers are obsolete. Traditional publishers offer important support to authors in the production, distribution and marketing of their works, without requiring any up-front expenditure by the author. Large publishing houses have in-house expertise at each of these different stages. Most self-published authors will struggle to fill all these roles to such a high standard themselves, and so may need to contract this work out to experts if they want to compete in a potentially global marketplace. Overall, self-publishing requires authors to be real entrepreneurs.

⁶³ See extended abstract by Raffaelli, R. "Reframing Collective Identity in Response to Multiple Technological Discontinuities: The Novel Resurgence of Independent Booksellers, 1995-2015" (3½ year field data collection effort complete, first manuscript in preparation).

Ultimately, it may not be the power of internet platforms which shape this industry, other than possibly for the narrow set of best-sellers, but rather the ways in which authors and consumers deal with the ‘long-tail’ problem. The risk of market failure has thus shifted focus to the implication of asymmetric information about the quality of books and future interventions may be necessary to ensure that the market functions in the interest of both readers and authors.

Although the paper has focused on the US and UK market, similar development are happening elsewhere. Clearly diffusion of existing digital innovation is taking place outside the US-UK markets. But, so are innovations as illustrated, for example, by the new French digital platform Librinova. One of the apparent barriers to innovation is the tradition in a number of countries, including EU member states, of having fixed prices for books, an intervention which, because it is carried out on a national basis, is beyond the reach of EU competition law enforcement (Poort and van Eijk, 2017). Smaller countries, such as Denmark, that have abandoned fixed book prices appear to perform well on most relevant fronts, including the number of books produced and the proportion of the population that buys and reads books.

Acknowledgements

Hviid, Izquierdo Sanchez, and Jacques are all members of CREATE and financial support from CREATE is gratefully acknowledged. An earlier version of this paper was presented at the 2016 FSR C&M and CMPF Scientific Seminar on Competition and Regulation in Infrastructure and Digital Markets, Florence. We are grateful to the participants and, in particular, to our discussant, Virginia Silvestri, for helpful comments and suggestions. The authors would also like to thank the editor and an anonymous referee from this journal, Lynne Chave, David Reader, Ruth Towse and an anonymous referee for the CREATE working paper series for their comments on an earlier draft. The usual disclaimer applies.

References

- Anderson, M. and Madruger, J. (2012). "Learning from the Crowd: Regression Discontinuity Estimates of the Effects of an Online Review Database", *Economic Journal*, 122 (563): 957-989.
- Baye, M. R., De Los Santos, B., and Wildenbeest, M. R. (2013). "Searching for Physical and Digital Media: the Evolution of Platforms for Finding Books", *NBER Chapters*, in: *Economic Analysis of the Digital Economy*: 137-165. National Bureau of Economic Research.
- Beck, J. (2012). "Advance Contracting, World-of-Mouth, and New-Product Success in Creative Industries: a Quantification for Books", *Journal of Media Economics*, 25(2): 75-97.
- Becker, B., Buhse, W., Günnewig, D., and Rump, N. (2003). "Digital Rights Management: Technological, Economic, Legal and Political Aspects", *Springer*.
- Brynjolfsson, E., Hu, Y. J., and Smith, M. D. (2006). "From Niches to Riches: Anatomy of the Long Tail", *Sloan Management Review*, 47(4): 67-71.
- Brynjolfsson, E., Hu, Y. J., and Smith, M. D., 2010, "Research Commentary-Long Tails vs. Superstars: the Effect of Information Technology on Product Variety and Sales Concentration Patterns", *Information Systems Research*, 21 (4): 736-747
- Buydens, M., and Dusollier, S. (2001). "Les Exceptions au Droit D'auteur: Evolutions Dangereuses", *22 CCE 10*.
- Carolan, S. and Evain, C. (2013), "Self-Publishing: Opportunities and Threats in a New Age of Mass Culture", *Publishing Research Quarterly*, 29(4): 285-300.
- Chevalier, J.A., Beinecke, W.S., and Mayzlin, D. (2006). "The effect of work of mouth on sales: Online book reviews", *Journal of Marketing Research*, 43(3): 345-354.
- D'Agostino, G. (2010). "Copyright, Contracts, Creators New Media, New Rules", *Edward Elgar*.
- Galopin, B. (2012). "Les Exceptions à Usage Public en Droit D'Auteur", *IRPI*.
- Geiger, C. (2004). "Droit D'Auteur et Droit du Public à L'Information", *IRPI*.
- Gilbert, R. J. (2015). "E-books: a Tale of Digital Disruption", *Journal of Economic Perspectives*, 29(3): 165-184.
- Goldstein, P. (2003). "Copyright's Highway: From Gutenberg to the Celestial Jukebox", *Stanford Law and Politics*, *Stanford University Press*.
- Greco, A. N. (1945). "The Book Publishing Industry", *Boston: Alyn & Bacon, 2nd Edition*.
- Grosheide, WF. (2001). "Copyright Law from a User's Perspective: Access Rights for Users", *European Intellectual Property Review*, 23(7): 321-325.
- Guibault, L. (2002). "Copyright and Contracts: an Analysis of the Contractual Overridability of Limitations on Copyright", *Kluwer*.
- Hardy, W., Krawczyk, M., and Tyrowicz, J. (2014). "Internet Piracy and Book Sales: a Field Experiment". *University of Warsaw, Faculty of Economic Sciences*. Working Paper No. 2014-23.

- Latcovich, S., and Smith, H. (2001). "Pricing, Sunk Cost, and Market Structure Online: Evidence from Book Retailing", *Oxford Review of Economic Policy*, 17(2): 217-234.
- Kay, J. (2016). "Rent Seeking in the Digital Economy", Keynote speech to the CREATE festival, London. Slides available at: <https://drive.google.com/file/d/0B0CpNLGeVtGlcNjyZmhXNThGeEk/view> .
- Lee, Y. J. Tan, Y. and Hosanagar, K. (2015). "Do I follow My Friends or the Crowd? Information Cascades in Online Film Rating", *Management Science*, 61(9): 2241-2258.
- Nelson, P. (1970). "Information and Consumer Behaviour", *Journal of Political Economy*, 78(2): 311-329.
- OECD (2012), "E-books: Developments and Policy Considerations", *OECD Digital Economy Papers*, No. 208, OECD Publishing. <http://dx.doi.org/10.1787/5k912zxcg5svh-en>
- Peukert, C., and Reimers I. (2018). "Digital Disintermediation and Efficiency in the Market for Ideas", *CESIFO*, working paper 6880.
- Poort, J., and van Eijk, N. (2017). "Digital fixation: the law and economics of a fixed e-book price", *International Journal of Cultural Policy*, 23(4): 464-481.
- Reimers, I. (2016). "Can Private Copyright Protection Be Effective? Evidence from Book Publishing", *Journal of Law and Economics*, 59(2): 411-440.
- Shaver, L. (2014) "Copyright and Inequality", *Washington University Law Review*, 92: 117-168.
- Sterling, J.A.L. (2015). "Sterling on World Copyright Law", *Sweet & Maxwell* (4th ed).
- Thompson, J. B. (2012). "Merchants of Culture", *Plume* (2nd edition).
- Vaver, D. (2006), "Publishers and copyright: Rights without duties?", Version of paper presented at the Sixth International Publishers Association copyright symposium, Montreal, April 23-25 2006, on Publishing & The Public Interest, at the Panel on the public image of publishers and copyright, Working Paper No 24/2006.
- Waldfoegel, J., and Reimers, I. (2015). "Storming the Gatekeepers: Digital Disintermediation in the Market for Books", *Information Economics and Policy*, 31: 47-58.
- Wiseman, L. (2007), "Copyright and the Regulation of the Australian Publishing Industry", in Carter, D. & Galligan, A. (Eds.) *Making Books: Contemporary Australian Publishing*. St Lucia, Queensland, University of Queensland Press.
- Zegners, D. (2016a). "Voluntary Disclosure of Product Information: The Case of E-book Samples", Available at SSRN: <http://ssrn.com/abstract=2380216>
- Zegners, D. (2016b). "Building an Online Reputation with Free Content: Evidence from the E-book Market", Available at SSRN: <http://ssrn.com/abstract=2753635>

Table 1: Number of UK book publishers registered for VAT – March to March

Year		2003/04	2004/05	2005/06	2006/07	2007/08	2008/09	2009/10
No	of	2,270	2,270	2,275	2,320	2,610	2,510	2,360
Publishers								
Year		2010/11	2011/12	2012/13	2013/14	2014/15	2015/16	2016/17
No	of	2,200	2,115	2,060	2,160	2,270	2,255	2,270
Publishers								

Source: ONS: UK Business: Activity, Size & Location various years⁶⁴

⁶⁴ <http://www.ons.gov.uk/ons/rel/bus-register/uk-business/index.html>.

Table 2: UK 4-firm concentration ratio

Year	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014
CR ₄	47.3	52.9	47.2	51.5	49.7	49.1	48.9	49.1	47.1	47.5

Source: UK Publishing Groups - Consumer Sales 2005-2014, BA Reports Library 2015

Table 3: levels of success

Now: Future:	Traditional Traditional	Traditional Self	Self Traditional	Self Self
Average number of traditionally published books	30.0	8.2	0.8	1.3
Average number of self-published books	2.2	2.2	7.1	10.0
Self-reported average income: traditional	\$73,523	\$32,946	\$1,119	\$1,582
Self-reported average income: self	\$4,816	\$6,970	\$15,137	\$45,310
Number reporting that they are earning a full time living of writing	69	20	10	350
% reporting that they are earning a full time living of writing	43%	27%	15%	26%
Total sample	159	88	66	1366

Table 4: Response to the question: If traditionally published,⁶⁵ which of the following are you satisfied with?

Now: Future:	Traditional Traditional	Traditional Self	Self Traditional	Self Self
The editing I received	83%	65%	25%	6%
The promotional efforts I received	42%	8%	6%	1%
My sales	42%	12%	13%	0%
The price of my works	54%	14%	19%	2%
The print distribution I received	51%	31%	13%	2%
The e-book distribution I received	60%	16%	19%	2%
My relationship with my publisher	67%	26%	19%	2%
Total sample	156	86	16	260

⁶⁵ Excluding authors who report never to have used a traditional publisher. The sample excludes five authors who self-report as traditionally published but who apparently have never published a book traditionally.

Table 5: Answer to the question: If self-published,⁶⁶ which of the following are you satisfied with?

Now: Future:	Traditional Traditional	Traditional Self	Self Traditional	Self Self
The quality of my final product	72%	84%	84%	84%
How much time I spend promoting	28%	35%	21%	30%
The distribution of my work	32%	45%	25%	50%
My current sales	20%	27%	10%	37%
Total sample	79	55	63	1348

⁶⁶ Excluding authors who report never to have self-published. Note that for those who currently see themselves as self-published, the sample excludes 21 respondents who apparently have never self-published.

Table 6: If self-published,⁶⁷ what degree of editing do you employ and how do you source your cover art?

Now: Future:		Traditional Traditional	Traditional Self	Self Traditional	Self Self
Total sample		79	54	63	1348
Editing	Hired freelance Editor	49.4%	57.4%	46.0%	54.0%
	Friends and family	13.9%	11.1%	28.6%	19.1%
	Critique group / other authors	19.0%	22.2%	20.6%	15.9%
	None	13.9%	7.4%	4.8%	9.9%
Cover	Hires professional cover designer	57.0%	66.7%	63.5%	55.6%
	Creates own	35.4%	29.6%	23.8%	39.4%
	Purchases pre-made covers	3.8%	3.7%	12.7%	4.7%

⁶⁷ We interpret this as respondents reporting that they have self-published at least one book.

Figure 1: Traditional structure

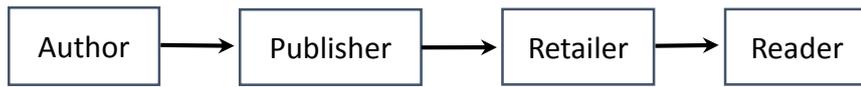
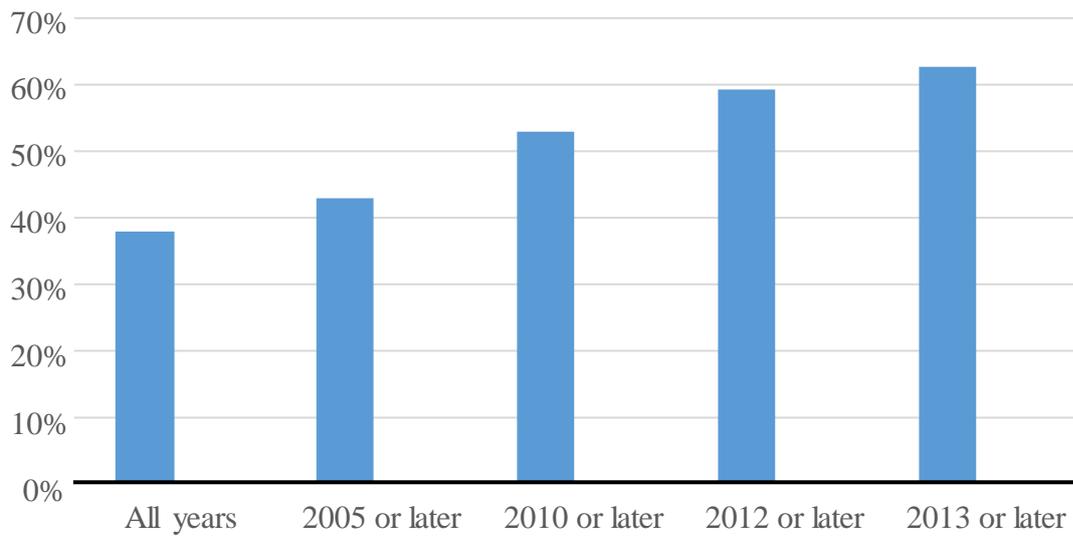


Figure 2: Share of self-published authors making at least \$10k per year



Source: *Author Earnings* Report: “Individual author earnings tracked across 7 quarters, Feb. 2014 – Sept. 2015”, available at: <http://authorearnings.com/report/individual-author-earnings-tracked-across-7-quarters-feb-2014-sept-2015/>.