CrossMark

Does the long tail really favor small publishers?

Stéphanie Peltier[1,2,3] · **Françoise Benhamou**[4] ·
Mamoudou Touré[4,5]

Abstract A growing body of literature is devoted to testing the reality of the
"long-tail" phenomenon. This literature is mostly, if not exclusively, focused on the
impact of Internet on the distribution of sales by product. However, the long tail also
raises the issue of a possible change in the usual market structure of cultural
industries: an oligopoly with a competitive fringe. To our knowledge, no paper
addresses the following question: If the long-tail effect does exist, is it of more
benefit to small or dominant publishers? The aim of this paper was to address this
issue in the context of the French publishing industry. Our main findings are as
follows: (1) the market concentration of the French book industry is lower online
than offline and (2) the difference in concentration between the two channels of
distribution tended to widen over the period 2004–2010. Strategies adopted by
leading publishers on the Web do not allow them to maintain the market share
obtained with bricks-and-mortar retailers. Furthermore, we show that the market

✉ Stéphanie Peltier
stephanie.peltier@univ-lr.fr

Françoise Benhamou
francoise.benhamou@univ-paris13.fr

Mamoudou Touré
toure@unistra.fr; toure@cournot.u-strasbg.fr

[1] LR-MOS CEREGE, University of La Rochelle, La Rochelle, France

[2] IRCAV, University of Sorbonne Nouvelle, Paris, France

[3] Faculté de Lettres, Langues, Arts et Sciences Humaines, Département LEA - 1, parvis Fernand
Braudel, 17 042 La Rochelle Cedex 01, France

[4] CEPN, University of Paris 13, 99, av. Jean-Baptiste Clément, 93430 Villetaneuse, France

[5] Pôle Européen de Gestion et d'Économie - PEGE, BETA University of Strasbourg,
61 Avenue de la Forêt Noire, 67085 Strasbourg Cedex, France

share lost by dominant firms is captured by small publishers online and by medium-sized publishers in conventional stores.

1 Introduction

Cultural industries usually share two common features (Caves 2000): an oligopolistic market structure with a competitive fringe, and sales highly concentrated on a small number of products, mainly because of the superstar effect (Rosen 1981) or "winner-take-all" phenomenon (Frank and Cook 1995). According to the long-tail effect (Anderson 2004, 2006), digitization in production, distribution, promotion and consumption should reduce the sales concentration for cultural products. But does digitization also impact the market structure of cultural industries? The impact of e-commerce on the distribution of sales has been the subject of much debate, but its effect on the strength of competition among publishers has been neglected. The present paper contributes to filling this gap in the literature.

According to Anderson (2004, 2006), the superstar effect tends to be offset by a "long-tail" effect in the digital era. Three forces tend to shift demand from the most popular products (the head of the sales distribution) to niche products (the tail). (1) The decrease in production costs increases the variety supplied and thus increases the length of the tail. (2) The constraints of physical shelf space disappear and distribution costs decrease drastically. Therefore, consumers have easier access to niche products. This fattens the tail because some products will find an online audience sufficient to ensure their profitability. (3) Furthermore, new ways of connecting demand and supply through Web 2.0 (Facebook, Twitter, blogs, forums, recommendation tools, etc.) improve matching between supply and demand by reducing search costs. Consumers may have better knowledge of products closer to their "ideal variety" than those that are highly promoted in the traditional media. Consequently, demand should switch from the mainstream products at the head of the demand curve (the hits) toward a huge number of niche products in the tail.

This prediction is widely debated, and both theoretical[1] and empirical works provide conflicting evidence about the existence and the magnitude of the long

---

[1] Brynjolfsson et al. (2006) or Hervas-Drane (2010) provides a theoretical framework that supports the long-tail effect. Conversely, some argue that the Internet reinforces the audience of best-selling products and thus sales concentration. This second category includes works that follow the "overchoice" theory (see, for instance, Gourville and Soman 2005), the pioneering works of Simon (1971) on the economics of attention, works that highlight a possible increase in marketing budget for a reduced number of superstars (Brynjolfsson et al. 2011) and works suggesting that some search engines and recommender tools (e.g., collaborative filters) based on sales and ratings can also reduce sales diversity (Fleder and Hosanagar 2009; Mooney and Roy 2000).

tail (see Table 1 for a survey of the empirical literature). As far as the publishing industry is concerned, Brynjolfsson et al. (2003, 2010), studying US book sales on Amazon, conclude that the concentration of sales did indeed decrease over the 2000s. In France, Bounie et al. (2010) found that the distribution of sales is less concentrated online than offline. Conversely, Benghozi and Benhamou (2010), working on a sample of French book sales, found that the long-tail effect remains very small online. However, these studies are based on rather limited samples (Benghozi and Benhamou 2010), on bestsellers only (Bounie et al. 2010) or on a mere estimation of sales (Brynjolfsson et al. 2003, 2010). Peltier and Moreau (2012) use comprehensive data over the period 2003–2007 and distinguish online and offline sales. They show that a long-tail effect exists in the French publishing industry with consumers shifting somewhat from bestsellers to medium or low sellers.[2] Farchy et al. (2013) provide an overview of the various impacts of digitization on the book industry: categories of works available, changes in consumer uses and cultural diversity. However, in all the papers above, a question is left unanswered: how digitization affects the market structure of the book industry?[3] Does it favor small publishers from the fringe or dominant firms? Put in other words, what is the profile of the firms whose product sales grow when sales increase on Internet distribution channels? These questions constitute our original contribution.

More precisely, we study which publishers benefit from the long-tail effect. Is this effect concentrated on small independent publishers (possibly specialized in niche products) or on big publishers and their subsidiary firms, or does it affect all kinds of publishing houses? The answer is far from straightforward. E-commerce might indeed favor limited audience books, but only those produced by large publishers. This would be the case if e-commerce generates new types of fixed costs that only large firms can afford. Do notice, however, that while individual publishers' strategies are key determinants of our results, we do not observe them and focus on their aggregate effect on the publishing industry market structure.

The paper is organized as follows: Section 2 introduces our research questions. Section 3 describes our empirical methodology. The results concerning the effect of the long tail are presented in Sect. 4. Section 5 briefly concludes by recalling the main findings and opening avenues for future research.

---

[2] The present paper relies on the same data source (GfK) as in Peltier and Moreau (2012) but with a radical difference. The latter paper deals with titles, whereas we look at publishers.

[3] Two papers also deal with the issue of the long tail at the firm level, but from a different perspective: Elberse (2008) and Bourreau et al. (2013). Elberse (2008) surveys different analyses on movies and music (especially from Rhapsody's transaction record and Quickflix movie transactions). She observes that the lengthening of the tail is very flat and derives managerial implications for producers. In particular, she concludes that the companies that will prosper are "the ones most capable of capitalizing on individual best sellers." Bourreau et al. (2013), on a sample of 151 French record companies, show that the companies that have adapted to digitization release more new albums without having higher overall sales. They conclude that these results are consistent with the long-tail hypothesis.

Table 1 Survey of the main empirical studies on the long-tail effect

Authors	Cultural industry	Country/firm	Period	Distinction online/offline	Data
Validation of the long-tail effect					
Brynjolfsson et al. (2003)	Book	Amazon US	2000	No	Estimation of sales
Brynjolfsson et al. (2010)	Book	Amazon US	Years 2000 and 2008	No	Estimation of sales
Bounie et al. (2010)	Book	France	22 weeks 2006	Yes	Top 100
Benghozi (2008)	Video DVD	France	2005	Yes	Comprehensive sales
Kumar et al. (2011)	Video	Amazon US	January 2008–March 2010	No	314 movies
Peltier and Moreau (2012)	Book	France	2003–2007	Yes	Comprehensive sales
Oestreicher-Singer and Sundararajan (2006)	Book	Amazon US	August 2005–na	No	250,000 books
Observation of a twofold phenomenon: long-tail and superstar effect and/or limited impact					
Elberse and Oberholzer-Gee (2008)	Video	USA	2000–2005	No	5500 titles
Benghozi and Benhamou (2010)	Music	France	2001–2005	Yes	Comprehensive sales
Benghozi and Benhamou (2010)	Book (novels)	France	2005 4 months (September–December)	Yes	693 titles
Elberse (2008)	Music	USA	January 2005 –April 2007	Yes	3300 artists
Reinforcement only of the Superstar effect					
Cha et al. (2007)	UGC	USA, Korea YouTube and Daum UGC	December 2006–March 2007	No	na
Tan and Netessine (2009)	Online video Web site	USA (ranking of movies on Netflix Web site)	2000–2005	No	100 millions of online rankings of 17,770 movies
Cho and Roy (2004)	UGC Page Rank Google	na	na	No	154 Web sites

Table 1 continued

Authors	Cultural industry	Country/firm	Period	Distinction online/offline	Data
Dellarocas and Narayan (2007)	Cinema and online video sales	Yahoo! Movies	2002	Yes	104 movies, 63,889 comments of 46,294 Internet users
Beuscart and Couronné (2009)	Music UGC	MySpace	November 2008	No	517 musicians

na Not available, *UGC* user-generated content

2 Research questions

Over the period 2004–2010, the French book market remained a duopoly with a double fringe.[4] Two companies dominated the industry: Hachette and Editis. In 2010, their respective turnovers were 2165 and 751 million euros. On the French market, their turnovers were about the same; the difference was mostly due to the very strong position of Hachette in foreign countries (especially the USA with its subsidiary Grand Central Publishing acquired from Time Warner in 2006). The first fringe comprised medium-sized groups (in particular, La Martinière, Gallimard, Flammarion, Albin Michel). The second fringe comprised small and very small publishers.[5] Did the development of IT over the period allow medium-sized and small companies to benefit from a long-tail effect by increasing their market share?[6] To address this issue, we study whether the long-tail phenomenon observed in the French book industry (Peltier and Moreau 2012) favors the "competitive fringe" of publishers. We therefore pose two research questions.

2.1 R1: Is the concentration of the book market weaker online than offline?

Digitization may lead to a lower concentration of the market through three effects. First, Internet favors the entry of new publishers into the market by reducing distribution costs. Second, online methods of marketing (Facebook, blogs, recommendation tools, etc.) are more open to small publishers. Although it is costly to reach the top of Google search ranking, the possibility of reaching consumers and capturing their attention is higher. Facebook pages, twits, etc. allow the production of information and the building of reputation, generating network effects. Third, by improving the match between supply and demand, recommendation systems and online word of mouth should favor small publishers who often offer niche products, look for new talent and try to identify "gaps" in the supply of big publishers.

2.2 R2: Does the difference between online and offline concentration disappear over the period?

Two scenarios are equally plausible for the evolution of the difference between online and offline market concentration. First, big companies may succeed in adapting their supply over time. They improve their promotional methods on Internet, join social networks, improve their ranking on Google search results by

---

[4] In 2012, Gallimard purchased Flammarion. The new firm is now the third member of the oligopoly that dominates the French publishing industry.

[5] For more information on the publishing structure in France, see, among others, Legendre (2012), Rouet (2013), Moreau and Peltier (2015).

[6] This hierarchy can be marginally disturbed by a best-seller effect (i.e., a best seller published by an independent publisher who suddenly gains some market share).

buying "clicks," etc. They also produce more and more niche products. Moreover, they may buy small pure-player firms with skills in digitization. In this way, they can capture a large part of the long-tail effect. The second possibility depends on the ability of newcomers and independent publishers to defend their relative advantage with specialized or risk-taker readers and writers. In this case, big companies leave the long tail to independent publishers and try to increase the best seller (winner-take-all) effect, in which case the difference in concentration may increase over time. Alternatively, even if big companies succeed in using Internet to promote their low-seller books small publishers could prove more efficient in achieving this task. Our analysis produces some evidence to evaluate the respective likelihoods of these two scenarios.

3 Empirical methodology

3.1 Data

To capture the effect of the long-tail phenomenon on the relative market shares of different publishers,[7] we use a comprehensive database of annual sales of physical books by publisher over a period of seven years (2004–2010) obtained from the French subsidiary of the GfK group, one of the world's leading market research organizations. GfK tracks all book sales in almost all outlets in France.[8]

In 2010, GfK's panel included more than 3500 offline and online shops.[9] Although the number of shops taken into account significantly increased over the period, the extrapolation method used by GfK ensures the representativeness of the panel at the national level. Data provided by GfK focus on two genres: comic books and literature. Literature (including novels, poetry and nonfiction) is the leading segment of the French book market, accounting for 25 % of units sold in 2010 (SNE 2011), and is usually considered the most emblematic genre of the book industry. Comic books represent a smaller market with around 9 % of total units sold in 2010.

Within the database provided by GfK, data can be broken down by channel of distribution. This allows comparisons to be made between online (Amazon …) and traditional sales channels (bookshops, large stores specialized in cultural products,

---

[7] In this paper, publishers are defined as publishing houses. Sales from all imprints or subsidiaries of a given publishing house are thus gathered.

[8] We use the same data source as in Peltier and Moreau (2012). However, there are two main differences. We work on a larger time period, and, above all, data refer to sales by publisher instead of sales by title.

[9] According to GfK, all online French booksellers or French subsidiary of foreign booksellers are included in the dataset all over the period: the two dominant firms (Amazon.fr and Fnac.com) as well as smaller online booksellers (Chapitre.com …).

supermarkets). Data on e-books are not reported, but they still remain marginal in France. Digital book sales represented 1.8 % of the book market in 2010 (SNE 2011). Furthermore, to avoid the risk of including the same title twice in the database—and thus overestimating the number of books sold—we have excluded paperbacks.[10] Since GfK could not distinguish between books whose first edition was in paperback and books that were only reprinted, the former are also excluded from the database.

Our database contains more than 170,000 different titles published by about 4000 publishers over the whole period. In total, 78.4 % of these different titles were literature books and 21.6 % were comics. This yields a sample of more than 400 million copies sold. Online sales, in units, rose from 1.8 % of overall sales in 2004 to 6.6 % in 2010. Data provided by GfK allow us to know accurately how many copies of each of the 170,000 books have been sold each year in the two distribution channels (offline and online). In this paper, book sales are analyzed at the publisher level. Thus, we have gathered the annual sales of copies of all books released by a given publisher in a given genre and in a given distribution channel (offline or online). Tables in the "Appendix" present the main descriptive statistics of the database used in this paper.

3.2 Methodology

To test the difference in sales distribution between the Internet and other channels, following Brynjolfsson et al. (2011), we estimate the Pareto curve for sales by publisher. The equation for Internet and offline data is the following:

$$\ln(Sales_j^i) = \beta_0^i + \beta_1^i \ln(SalesRank_j^i) + \varepsilon_j^i \tag{1}$$

where $Sales_j^i$ denotes the level of sales for each publisher j over the period in the distribution channel i (offline or online). $SalesRank_j^i$ is an ordinal ranking of the frequency of sales of each publisher j in the distribution channel i. In this setting (model 1), β_1^i measures how quickly the sales of a given publisher in a channel decrease as the sales rank rises. The more strongly negative β_1^i is, the higher the market concentration.

If Internet sales of books by publisher are less concentrated than offline sales, we would expect β_1^i to be less strongly negative (i.e., lower in absolute value) in the Internet channel than in the conventional channel. This reflects the idea that low-selling publishers (i.e., publishers that obtain higher ranks[11]) obtain a larger share of sales in this channel. To test whether the β_1^i coefficient is significantly less negative for the Internet channel than for the physical channel, we pool Internet and offline data into one dataset (Brynjolfsson et al. 2011). Thus, the linear regression we estimate is the following:

---

[10] Furthermore, paperback books belonging by nature to the long tail, their exclusion and our focus on hardback books could only reinforce the robustness of our results.

[11] The first (lowest) rank corresponds to the publisher with the highest sales.

$$\ln(Sales_j^i) = \beta_0^i + \beta_1^i \ln(SalesRank_j^i) + \beta_2^i Internet_j + \beta_3^i Internet_j$$
$$\times \ln(SalesRank_j^i) + \varepsilon_j^i \qquad (2)$$

Internet is a dummy that indicates whether an observation is for the Internet and we introduce the interaction term between the variable *Internet* with ln(*SalesRank*). A positive value for β_3^i would indicate that the market concentration is lower online than offline.

However, in our database, this lower concentration of sales online may be a pure artifact. It could be explained by several biases. A first bias is related to the segment of books considered (comics or literature books). If the sales of one segment are less concentrated than those of the other segment and at the same time relatively much more important online than offline, online sales would appear less concentrated. We therefore include in our regression an interacted variable *SalesRank × Comics*, where *Comics* is a dummy variable equal to one if the distribution of sales refers to comics books and zero if it refers to literature books.

Likewise, a second bias could be due to temporal specificities over the period. We therefore introduce a continuous variable *Year* (ranging from 2004 to 2010) as well as the interaction *Year × SalesRank*.

We can also imagine that the number of titles available each year in both channels might mechanically affect the distribution of sales. For instance, if more different titles are sold on the Internet than offline, concentration by publisher could appear lower online. To check whether a long-tail effect is not only due to the fact that more references are directly available online, we include the variable *Titles* (which is the log of the number of titles sold per year for each genre and each channel of distribution) as well as the interaction *Titles × SalesRank*.

Another bias could be related to the specific life cycle a book usually experiences. When a lot of new titles are released during a year, sales are spread over a greater number of titles. If consumers buy more new titles when they use the Internet channel, the concentration of sales by publisher will mechanically be lower in this channel. The interaction *News × SalesRank*, where the variable *News* equals the log of the number of new titles that have been released during the same year, allows us to control for this bias.

The difference between online and offline sales concentration might also be explained by the prices charged for books in each channel. In the French case, this bias seems unlikely, because of the "fixed price agreement." The price of a book is decided by the publisher and is uniform for all retailers. However, as retailers can grant a 5 % discount, we introduce the mean price[12] for each

---

[12] The price for a given book, in a given channel, and for a given time period is calculated as the ratio "sales"/"units sold." Thus, it is the average actual price that is taken into account rather than the listed price.

Table 2 Evolution of the number of publishers by genre and channel of distribution (2004–2010)

	2004	2005	2006	2007	2008	2009	2010	Variation 2004–2010
Comics								
Internet	352	419	455	523	624	723	735	+383
Offline	542	574	602	643	662	767	818	+276
Literature								
Internet	1.512	1.784	1.978	2.224	2.601	2.962	3.101	+1.589
Offline	2.236	2.384	2.602	2.637	2.781	3.046	3.264	+1.028

publisher (*Price*) and the interacted variable *SalesRank* × *Price* in our regressions.

The long-tail effect may also result from a higher increase in the number of publishers online than offline over the period, which would naturally lead to a lower concentration of the market on the Internet channel. As Table 2 shows, the number of publishers did indeed rise faster online than offline, for both literature and comic books.

To check whether the long-tail effect is not only due to a higher increase in the number of publishers who sold online than offline, we also introduce a control variable *SalesRank* × *Publishers*, where the variable *Publishers* is the log of the number of publishers who have sold at least one copy (for a given year, a given channel and a given genre).

To assess whether the long-tail effect is not just a temporary phenomenon, we add to the above model the "*Internet* × *SalesRank* × *Year*" variable (model 3). To answer our second research question, "*Does the difference between online and offline concentration disappear over the period?*", we observe the coefficient for this interaction variable. If it is negative, we conclude that the difference in concentration vanishes over the period.

A drawback of the above models is that they do not shed any light on the evolution of the market shares of the various types of firms presented above: duopoly, first fringe and second fringe. A lower concentration could be due to increased market share of the smallest publishers (second fringe), a rise in the medium-sized publishers (first fringe) at the expense of the dominant firms (duopoly) or to a mere reallocation of market shares between these duopolistic firms. To address this issue, we study the market share of the 2, 4, 10, 20 and 50 biggest publishers (CR2, CR4, CR10, CR20 and CR50) for each distribution channel (online vs. offline) and for each segment (literature vs. comics).

Table 3 Pareto curve estimates in value [14 subsamples pooled for both Internet sales and offline sales—publishers (LR)]

	Internet	Offline
Constant	15.894***	21.038***
	(0.050)	(0.060)
Sales rank	−1.610***	−2.106***
	(0.007)	(0.009)
Adjusted R^2	0.690	0.697
Sample size	19.991	23.463

Standard errors are in brackets

LR linear regression

*** Significance at 1 %. All the results are confirmed by the quantile regressions (not reproduce here)

4 Results

To conduct our analysis, we construct twenty-eight subsets pooled in our data according to genre (comic books or literature), the distribution channel (online or offline) and a given year of the 7-year period: 7 years × 2 genres × 2 distribution channels.

4.1 Is the concentration of the book market weaker online than offline?

We first study the difference in market concentration between both online and offline channels without any control variables (Table 3).

Results suggest that sales by publisher decrease more slowly as the rank increases in the Internet than in conventional stores. Model 1 in Table 4 provides a test of the significance of this result. The coefficient of *SalesRank* × *Internet* is, as expected, positive and highly significant at the 1 % level. Our first hypothesis is thus supported: The market concentration is lower online than offline.

Model 2 in Table 4 provides a robustness test of this result by controlling for the diverse variables (included the number of publishers) that could impact the online market concentration. The coefficient of *SalesRank* × *Internet* remains positive and highly significant at the 1 % level.

4.2 Does the difference between online and offline concentration disappear over the period?

However, the difference in market concentration between online and offline sales may be temporary. Over time, dominant publishers might have adapted their strategies to improve their market share on the Web. Likewise, on the demand side, the predominance of early adopters of e-commerce with niche tastes that small

Table 4 Pareto curve estimates in value [28 subsamples pooled—publishers (LR)]

	Model 1	Model 2	Model 3
Constant	21.208***	−546.065***	−829.167***
	(0.084)	(110.277)	(109.378)
Sales rank	−2.150***	138.604***	181.582***
	(0.012)	(17.702)	(17.559)
Internet	−3.470***	−3.786***	1370.248***
	(0.100)	(0.070)	(66.085)
Price		−11.162***	−7.863***
		(1.060)	(1.054)
Titles		1.030*	3.584***
		(0.565)	(0.593)
Publishers		−5.229***	−2.790***
		(0.819)	(0.814)
News		5.854***	5.075***
		(0.245)	(0.244)
Comics		−6.298***	0.721
		(1.195)	(1.220)
Year		0.289***	0.406***
		(0.058)	(0.057)
Sales Rank × Internet	0.340***	0.576***	−162.713***
	(0.014)	(0.010)	(10.304)
Sales Rank × Price		1.905***	1.261***
		(0.176)	(0.175)
Sales Rank × Titles		−1.254***	−1.596***
		(0.085)	(0.091)
Sales Rank × Publishers		2.282***	2.044***
		(0.133)	(0.132)
Sales Rank × News		−0.702***	−0.595***
		(0.039)	(0.038)
Sales Rank × Comic		1.621***	0.716***
		(0.198)	(0.201)
Sales Rank × Year		−0.072***	−0.090***
		(0.009)	(0.009)
Internet × Year			−0.684***
			(0.032)
Internet × Sales Rank × Year			0.081***
			(0.005)
Adjusted R^2	0.642	0.897	0.899
Sample size	43,454	43,454	43,454

Standard errors are in brackets

LR linear regression

*** Significance at 1 %. All the results are confirmed by the quantile regressions (not reproduce here)

Table 5 Concentration ratio online and offline, literature and comics, 2004–2010

	2004 (%)	2005 (%)	2006 (%)	2007 (%)	2008 (%)	2009 (%)	2010 (%)	Variation 2004–2010 (in points of %)
Online								
Literature								
CR2	36.99	41.10	35.50	33.67	32.15	32.70	30.36	−6.63
CR4–CR2	22.42	18.98	21.64	21.99	19.71	16.75	17.08	−5.34
CR10–CR4	16.22	17.08	18.15	19.52	21.50	21.45	22.69	+6.47
CR20–CR10	8.28	7.07	6.93	6.38	7.48	6.79	6.52	−1.76
CR50–CR20	6.62	5.99	6.74	6.77	6.27	6.80	7.04	+0.42
100–CR50	9.47	9.78	11.04	11.67	12.89	15.51	16.31	+6.84
Comics								
CR2	48.91	45.78	44.59	44.32	43.36	40.67	42.46	−6.45
CR4–CR2	22.49	20.60	21.78	21.51	21.83	22.12	20.17	−2.32
CR10–CR4	16.46	19.70	20.81	19.69	20.98	21.72	21.38	+4.92
CR20–CR10	5.80	6.44	5.71	6.41	5.47	6.43	6.30	+0.50
CR50–CR20	4.20	5.04	4.65	5.13	4.77	5.16	5.49	+1.29
100–CR50	2.14	2.44	2.46	2.94	3.59	3.90	4.20	+2.06
Offline								
Literature								
CR2	45.42	49.79	43.72	39.61	38.26	40.28	37.51	−7.91
CR4–CR2	23.48	18.24	21.63	23.98	21.06	19.53	18.98	−4.50
CR10–CR4	15.07	16.52	18.02	18.51	22.75	21.68	24.28	+9.21
CR20–CR10	8.04	6.36	5.89	5.94	6.25	5.78	6.14	−1.90
CR50–CR20	4.57	4.82	5.70	6.44	5.95	6.37	6.73	+2.16
100–CR50	3.42	4.27	5.04	5.52	5.73	6.36	6.36	+2.94
Comics								
CR2	54.43	48.33	46.89	44.00	43.66	42.03	43.42	−11.01
CR4–CR2	21.95	21.23	23.01	22.21	22.47	20.09	20.45	−1.50
CR10–CR4	15.74	21.03	19.07	19.62	22.24	25.77	23.13	+7.39
CR20–CR10	4.50	4.77	6.20	8.15	5.62	5.88	5.91	+1.41
CR50–CR20	2.59	3.38	3.39	4.11	3.88	4.30	4.85	+2.26
100–CR50	0.79	1.26	1.44	1.91	2.13	1.93	2.24	+1.45

publishers supply better might vanish with the increase in the number of consumers who purchase books online.

Model 3 in Table 4, which introduces the interacted variable *Internet × SalesRank × Year*, does not support this scenario. Indeed, the sign of the coefficient is significantly positive. Thus, our results highlight an ever wider difference in concentration between online and offline sales over the period. This

result suggests that as yet, the dominant firms have not adapted their strategies to the online market sufficiently to maintain the market share they enjoyed in bricks-and-mortar retailers.

4.3 Is there a significant role for small publishers in the online market?

We have shown that market concentration in the book industry is lower online than offline, at least for the two segments studied. We have also found that this difference does not disappear over the period 2004–2010. Table 5 allows us to shed further light on the comparative level of market concentration online and offline and its evolution over the period.

It turns out that the lower market concentration observed online compared with offline is related to a loss in the market share of the duopolists and of firms at the top of the first fringe (up to 10th rank). On average, over the period and for both literature and comics, the top ten firms perform less well online than offline. The duopolists experience the biggest fall, while the gain in market share is obtained by second fringe firms (ranking above 50th). In keeping with the long-tail hypothesis, the smallest publishers seem to particularly benefit from the rise of the online market. It allows them to overcome the disadvantage of the limited space devoted to their books in conventional retail stores. In France, the distribution networks supplying booksellers belong to the largest publishers. Their bargaining power is thus much higher with conventional retailers than with online retailers such as Amazon. Moreover, the online market is probably less favorable to firms that rely on traditional marketing campaigns to promote their books. Conversely, small publishers, who usually promote less popular authors, are less disadvantaged when using online promotion and recommendation tools. However, it is interesting to note that the differences in concentration between online and offline markets are wider for the literature segment than for comics. On average over the period, in the literature segment, the online market share of the duopoly was 7.5 % below their offline market share, whereas the firms above the 50th rank had a total online market share 7 % above their offline level. In the comics segment, these two figures were −2 and +1.4 %, respectively.

Table 5 also allows us to better understand the dynamics at work over the period. Whatever the market segment and the channel of distribution, market concentration tended to decrease over the period 2004–2010. The four leading firms lost about 10 % of their market share, while two types of firms benefited from this weakening, depending on the channel of distribution and the segment considered. Both online and offline and both in literature and in comic books, the firms of the first fringe (between 5th and 10th rank) enjoyed growth in their market share between 2004 and 2010. For literature books, the gain was +6.5 % on the online market and the increase was 9.2 % on the offline. For comic books, these figures were +4.9 and +7.4 %, respectively. The second type of firms that benefited from the weakening of market leaders was the firms of the second fringe. But this result only holds

online and for literature books (+6.8 %). The impact of IT on market concentration in the comic segment is quite counterintuitive. Since comic book readers are usually younger and more familiar with digital technology, we would have expected the development of online distribution and recommendation tools to lead to a larger fall in concentration for comic books to the benefit of small publishers able to match young consumers' preferences. Our results show that consumers' tastes still focus on best sellers, both offline and online. The importance of mimetic behavior, reinforced by the broadcasting of TV programs that are derivative products of best-selling comic books, probably partially explains this result.

5 Conclusion

The paper contributes to the empirical studies on the long-tail effect by analyzing its consequences on the market structure of the book industry. Four main issues are highlighted. First, a long-tail effect exists when sales are counted by publisher. In the French book industry, the lesser concentration of sales online versus offline (Peltier and Moreau 2012) goes hand in hand with a deconcentration of the market structure. Second, this trend is more and more evident over the period 2004–2010. Third, the type of firms that benefit from the erosion of the leaders' market share, both offline and online, depends on the segment of publishing activity considered. In the comics segment, it is the first fringe of publishers that most benefit from the dominant firms' loss of market share online, while in the literature segment it is the smallest firms of the second fringe that most benefit. Finally, we show that the rise of online sales does not drastically change the relative level of concentration when winner-take-all habits are frequent (as in the case of comics).

Further research is needed on at least two issues. First, it would be interesting to investigate the specific role of independent publishers in the top 100 sales. This would show whether the increase in the market share of these publishers is due to a rise in the sales of long-tail books alone, or if the position of the firms at the core of the oligopoly is also undermined for best sellers. Second, it would be interesting to test the long-tail effect on the e-book market, which has developed enormously since 2007, especially in the USA.[13]

---

[13] In France, sales of digital books were estimated to be 0.6 % of total book sales (according to Gfk in 2012), but with more than 80 % of growth in 2012.

Appendix

See Tables 6 and 7.

Table 6 Descriptive statistics at the overall sample level (2004–2010)

Genre	Channel		Total
Comic	Online	# Titles	31,219
		# News titles	3915
		# Units sold	7,250,224
		Sales	87,712,552
	Offline	# Titles	35,468
		# News titles	5748
		# Units sold	177,841,248
		Sales	1,974,530,048
Literature	Online	# Titles	106,450
		# News titles	11,959
		# Units sold	8,732,049
		Sales	156,736,384
	Offline	# Titles	118,651
		# News titles	20,556
		# Units sold	207,181,728
		Sales	3,780,294,400
Comic	All	# Titles	36,841
		# News titles	6220
		# Units sold	187,948,064
		Sales	2,095,365,376
Literature	All	# Titles	133,651
		# News titles	26,531
		# Units sold	219,408,256
		Sales	4,003,258,624
All	Online	# Titles	137,661
		# News titles	15,874
		# Units sold	15,982,273
		Sales	244,448,928
	Offline	# Titles	154,110
		# News titles	26,304
		# Units sold	385,022,912
		Sales	5,754,823,168
All	All	# Titles	170,492
		# News titles	32,751
		# Units sold	407,356,320
		Sales	6,098,624,000

Sales are in euros

Table 7 Yearly descriptive statistics

Genre	Channel		2004	2005	2006	2007	2008	2009	2010
Comics	Online	# Titles	8007	10,061	11,188	13,512	16,026	18,456	19,204
		# News titles	374	357	474	542	670	866	632
		# Units sold	539,325	637,509	740,250	994,709	1,182,000	1,422,656	1,733,775
		Sales	6,177,981	7,269,491	8,772,252	11,765,329	14,289,232	17,587,284	21,850,980
	Offline	# Titles	14,281	15,380	16,722	17,630	18,445	20,942	23,729
		# News titles	733	712	790	824	906	1001	782
		# Units sold	31,398,766	28,745,332	25,425,444	24,165,220	21,460,056	23,224,392	23,422,040
		Sales	325,632,160	303,803,584	277,335,392	268,482,848	244,641,088	271,326,688	283,308,224
Literature	Online	# Titles	20,871	26,754	29,987	36,195	47,465	55,867	56,768
		# News titles	754	1049	1156	1402	2404	3061	2133
		# Units sold	619,814	727,012	925,734	1,151,066	1,527,466	1,774,260	2,006,697
		Sales	11,147,715	13,273,928	16,785,292	20,751,614	28,066,652	31,754,606	34,956,572
	Offline	# Titles	45,742	46,070	48,409	47,712	44,948	48,274	50,895
		# News titles	2893	3116	3393	2796	3043	3054	2261
		# Units sold	29,447,642	29,353,992	29,910,948	29,202,876	29,026,758	30,948,594	29,290,912
		Sales	528,505,962	529,872,896	538,653,120	527,562,208	540,985,536	574,781,888	539,932,800
Comics	All	# Titles	15,937	17,441	19,019	21,132	23,486	26,497	28,764
		# News titles	769	750	837	901	979	1158	826
		# Units sold	32,194,384	29,473,320	26,442,768	25,767,254	23,911,598	24,803,088	25,355,652
		Sales	334,866,752	312,156,032	289,686,048	287,034,912	272,648,384	291,015,168	307,958,144
Literature	All	# Titles	56,650	60,532	64,301	68,650	76,452	84,977	87,837
		# News titles	3129	3542	3736	3358	4415	4861	3490
		# Units sold	30,284,720	30,302,040	31,123,096	31,354,474	31,460,998	33,141,192	31,741,732
		Sales	543,470,272	547,011,392	560,463,680	568,751,424	586,633,216	613,994,240	582,934,336

Table 7 continued

Genre	Channel		2004	2005	2006	2007	2008	2009	2010
All	Online	# Titles	28,878	36,815	41,175	49,707	63,491	74,323	75,972
		# News titles	1128	1406	1630	1944	3074	3927	2765
		# Units sold	1,159,139	1,364,521	1,665,984	2,145,775	2,709,466	3,196,916	3,740,472
		Sales	17,325,696	20,543,420	25,557,544	32,516,944	42,355,884	49,341,888	56,807,552
	Offline	# Titles	60,023	61,450	65,131	65,342	63,393	69,216	74,624
		# News titles	3626	3828	4183	3620	3949	4055	3043
		# Units sold	60,846,408	58,099,324	55,336,392	53,368,096	50,486,816	54,172,984	52,712,952
		Sales	854,138,112	833,676,480	815,988,480	796,045,056	785,626,624	846,108,608	823,241,024
All	All	# Titles	72,587	77,973	83,320	89,782	99,938	111,474	116,601
		# News titles	3898	4292	4573	4259	5253	6019	4316
		# Units sold	62,479,104	59,775,360	57,565,864	57,121,728	55,372,596	57,944,280	57,097,384
		Sales	878,337,024	859,167,424	850,149,696	855,786,304	859,281,600	905,009,408	890,892,480

Sales are in euros

References

Anderson, C. (2004). The long tail. *Wired Magazine, 12*(10), 170–177.

Anderson, C. (2006). *The long tail*. New York: Hyperion press.

Benghozi, P. J. (Ed.) (2008). *Effet longue traîne ou effet podium: une analyse empirique des ventes de produits culturels en France*. Paris: Ecole Polytechnique.

Benghozi, P. J., & Benhamou, F. (2010). The long tail: Myth or reality? *International Journal of Art Management, 12*(3), 43–53.

Beuscart, J. S., & Couronné, T. (2009). La distribution de la notoriété artistique en ligne: une analyse quantitative de MySpace. *Terrains et Travaux, 15*, 147–170.

Bounie, D., Eang, B., & Waelbroeck, P. (2010). Marché Internet et réseaux physiques: comparaison des ventes de livres en France. *Revue d'Economie Politique, 120*(1), 141–162.

Bourreau, M., Gensollen, M., Moreau, F., & Waelbroeck, P. (2013). "Selling less of more?" The impact of digitization on record companies. *Journal of Cultural Economics, 37*(3), 327–346.

Brynjolfsson, E., Hu, Y. J., & Simester, D. (2011). Goodbye pareto principle, hello long tail: The effect of search costs on the concentration of product sales. *Management Science, 57*(8), 1373–1386.

Brynjolfsson, E., Hu, Y. J., & Smith, M. D. (2003). Consumer surplus in the digital economy: Estimating the value of increased product variety at online booksellers. *Management Science, 49*(11), 1580–1596.

Brynjolfsson, E., Hu, Y. J., & Smith, M. D. (2006). From niches to riches: The anatomy of the long tail. *Sloan Management Review, 47*(4), 67–71.

Brynjolfsson, E., Hu, Y. J., & Smith, M. D. (2010). The longer tail: The changing shape of Amazon's sales distribution curve. *Working Paper*. Available at http://ssrn.com/abstract=1679991. Accessed 30 June 2011.

Caves, R. E. (2000). *Creative industries, contracts between art and commerce*. Cambridge, MA: Harvard University Press.

Cha, M., Kwak, H., Rodriguez, P., Ahn, Y.-Y., & Moon, S. (2007). I tube, you tube, everybody tubes: Analysing the world's largest user generated content video system. In *IMC'07 conference*, San Diego.

Cho, J., & Roy, S. (2004). Impact of search engines on page popularity. In *WWW2004 conference*.

Dellarocas, C., & Narayan, R. (2007). Tall heads vs. long tails: Do consumer reviews increase the informational inequality between hit and niche products? In *Robert H. Smith School of Business Research Paper* No. 06-056. Available at SSRN: http://ssrn.com/abstract=1105956.

Elberse, A. (2008). Should you invest in the long tail? *Harvard Business Review, 86*(7/8), 88–96.

Elberse, A., & Oberholzer-Gee, F. (2008). Superstars and underdogs: An examination of the long tail phenomenon in video sales. In *Harvard Business School Working Paper* No. 07-015.

Farchy, J, Gansemer, M., & Petrou, J. (2013). eBook and book publishing, Chap. 31. In Towse & Handke (Eds.), Handbook on the digital creative economy. E. Elgar.

Fleder, D., & Hosanagar, K. (2009). Culture's next rise or fall: The impact of recommender systems on sales diversity. *Management Science, 55*(5), 697–712.

Frank, R., & Cook, P. (1995). *The winner-take-all society*. New York: The Free Press.

Gourville, J. T., & Soman, D. (2005). Overchoice and assortment type: When and why variety backfires. *Marketing Science, 24*(3), 382–395.

Hervas-Drane, A. (2010). *Word of mouth and taste matching: A theory of the long tail*. Mimeo: University of Pompeu-Fabra.

Kumar, A., Smith, M. D., & Telang, R. (2011). Information discovery and the long tail of motion picture content. In *WP, Carnegie Mellon University*, Pittsburgh, PA.

Legendre, B. (2012). Le Paysage éditorial. *Les Métiers de l'édition* (pp. 11–70). Paris: Editions du Cercle de la Librairie.

Mooney, R. J., & Roy, L. (2000). Content based book recommending using learning for text categorization. In *5th ACM conference on digital libraries* (pp. 195–204).

Moreau, F., & Peltier, S. (2015). *Fondamentaux et mutations du secteur de l'édition: les ressorts de l'économie de la création*. Paris: Syndicat National de l'Edition.

Oestreicher-Singer, G., & Sundararajan, A. (2006). Network structure and the long tail of electronic commerce. In *Proceedings of the 10th INFORMS conference on information systems and technology*.

Peltier, S., & Moreau, F. (2012). Internet and the "long tail versus superstar effect" debate: Evidence from the French book market. *Applied Economics Letters, 19*(8), 711–715.

Rosen, S. (1981). The economics of superstars. *American Economic Review, 71*(5), 845–858.

Rouet, F. (2013). *Le livre: une filière en danger* (4ème éd.). Paris: collection «les études», La Documentation française.

Simon, H. A. (1971). Designing organizations for an information-rich world. In M. Greenberger (Ed.), *Computers, communication, and the public interest*. Baltimore: The Johns Hopkins Press.

SNE (2011). *Les chiffres clés de l'édition 2011, données 2010*. Paris: Syndicat National de l'Edition.

Tan, T. F., & Netessine, S. (2009). Is Tom Cruise threatened? Using Netflix Prize data to examine the long tail of electronic commerce. In *Wharton Working Paper*.