



No. 12-2020

**Georg Götz, Daniel Herold, Phil-Andrian Klotz and Jan
Thomas Schäfer**

**Innovation, Bestsellers and Digitization – Where to Find the
Needle in the Haystack?**

This paper can be downloaded from
<http://www.uni-marburg.de/fb02/makro/forschung/magkspapers>

Coordination: Bernd Hayo • Philipps-University Marburg
School of Business and Economics • Universitätsstraße 24, D-35032 Marburg
Tel: +49-6421-2823091, Fax: +49-6421-2823088, e-mail: hayo@wiwi.uni-marburg.de

Innovation, Bestsellers and Digitization - Where to Find the Needle in the Haystack?

Georg Götz * Daniel Herold* Phil-Adrian Klotz*
Jan Thomas Schäfer*[†]

March 19, 2020

Abstract

We empirically analyze the role of e-Commerce and brick-and-mortar retailers in discovering new bestsellers in the book market, using the German market as an example. Using an AR(1)-Process, we find that when a title becomes a bestseller sales increase in e-Commerce and decrease in brick-and-mortar stores relative to a title that is not on the bestseller-list. This finding implies that consumers in the online channel respond by increasing sales upon receiving a quality signal (i.e., a title becoming a bestseller). Consumers in the offline channel seem to already know that title so that, compared to a title that is not on the bestseller-lists, sales are decreasing. This might imply that consumers in the offline channel are more likely to read future bestsellers.

1 Introduction

This paper analyzes the discovery of new products in different retail channels, namely e-Commerce and brick-and-mortar stores. We use data on the book market to analyze in which channel the adoption of innovation is faster, i.e., in which channel new bestsellers are purchased earlier, and which channel relies more strongly on external signals. This approach allows to identify which channel has a greater impact on the discovery of new bestsellers. We define bestsellers according to well-known lists like the The New York Times bestseller list, i.e. the success of a book is mainly measured based on sales figures.

The book market is well-suited to analyze the discovery of innovation. Although books can be seen as horizontally and vertically differentiated, each newly

*Chair for Industrial Organization, Regulation and Antitrust, Department of Economics, Justus Liebig University Giessen. Licher Strasse 62, 35394 Giessen, Germany.

[†]Corresponding author. E-mail: jan.t.schaefer@wirtschaft.uni-giessen.de.

published book can be seen as an innovation. It is particularly easy to measure the success of such a new product from a purely commercial perspective: If a book enters the bestseller lists this is a sign of the commercial success of that title.

In the last two decades, digitization has led to a sharp increase in new cultural products, including books (Waldfogel, 2015). The diffusion and discovery of innovative and commercially promising titles was always a challenging task, which, given the large number of new books, became even more difficult in the era of digitization. The process of discovery and the diffusion of new titles is described in Chapter 5 in Waldfogel (2018). From the viewpoint of an author, of course, the first step is to create something new and innovative, which meets the tastes of a large enough number of potential customers. If such a title is created, the next obstacle is to find a publisher, who exerts the proper effort in promoting and selling the title. These so-called *gatekeepers* can be avoided when the author chooses to become a publisher herself through self-publishing. With the increasing popularity of e-Books and the rise of e-Commerce, this option became more attractive in the recent past. This increase in attractiveness, however, led to a drastic increase in the range of titles but not necessarily to an increase in quality. Waldfogel (2018, p.133) visualizes the situation by the metaphor of “finding a needle in the haystack”: it is hard for consumers to find a high-quality title among a large number of (new) titles, and it is hard for an author to get found. Digitization most likely has led to more “needles”, however, at the same time the size of the “haystack” increased as well.

We analyze the role e-Commerce and brick-and-mortar stores in this process of new-product discovery. In particular, we investigate in which channel commercially successful titles are discovered earlier. In this context, both retail channels have certain (dis-)advantages, which will be discussed below. The criterion which we attribute a title to be commercially successful is whether the title appeared on a bestseller-list visible to a large number of consumers. In that respect, bestseller-lists can act as a vehicle to coordinate consumer-behavior (Banerjee, 1992) or as a signal of quality (Clement et al., 2007; Sorensen, 2007). For instance, Sorensen (2007) finds that sales increase if a title appears on popular bestseller-lists, which means that bestseller-lists act as a signal of quality for consumers. This concept of quality-signaling also relates to other cultural products such as movies (King, 2007; Reinstein & Snyder, 2005) when it comes to movie critiques or music when it comes to playlists on, e.g., spotify (Aguiar & Waldfogel, 2018). For the case of books, the likelihood that a title is considered a commercial success is particularly low. For instance, Beck (2012) estimates a success-rate for books of 10%-15%. This feature of relatively low success rates is also common to all cultural goods, including music and movies (Waldfogel, 2016).

In order to identify the impact of the different retail channels on the discovery of new products, we employ an approach comparable to Sorensen (2007), who uses an AR(1)-model to analyze the impact of bestseller-lists on sales. Using an AR(1)-model, we find that at the point an average title enters the bestseller-list sales from week $t - 1$ to week t (i) increase by approximately 13% in e-Commerce and (ii) decrease by approximately 19% at brick-and-mortar retailers compared to a title that is not on the bestseller-list. Given that sales are decreasing over time, this result implies that sales decrease approximately 13% slower in e-Commerce and approximately 19% faster in brick-and-mortar stores. This means that e-Commerce relies more strongly on an external quality signal, i.e., a title becoming a bestseller than it is the case at brick-and-mortar stores. In the offline channel, the proportion of sales of titles which are *not* bestsellers is therefore significantly higher. This means that *ceteris paribus* a large share of consumers purchase those titles offline which are either (i) future bestsellers, (ii) titles that will never become bestsellers or (iii) old bestsellers. The explanation (i) that consumers are more likely to purchase future bestsellers offline, or, in other words, that bestsellers are more likely to be “born” offline rather than online, seems realistic in a relatively large number of cases because the market share of brick-and-mortar stores is about three times the market share of e-Commerce (see Section 3 below). Hence, it could be that in a relatively large number of cases consumers start purchasing new bestsellers offline. When the title then becomes a bestseller, this signals the quality of that title. Upon receiving this signal, a significant number of consumers subsequently starts purchasing the title online. However, this has to be analyzed in further detail in future research.

Potential explanations for this finding are as follows. For instance, on the demand-side of the market, it could be that heavy book-buyers prefer to purchase books at brick-and-mortar retailers over e-Commerce. Given that these buyers read more, the odds that they find the “needle in the haystack” is higher. It could also be the case that consumers prefer to purchase online when they already know what they want to buy whereas when they not yet know which title to buy (or when it comes to impulse purchases) they prefer the offline channel. Given that innovative titles are inherently unknown at first consumers will not search for such titles in e-Commerce. A supply-side explanation of the finding is that in e-Commerce retailers start to actively promote a title only when the odds of economic success are high enough. Upon receiving a signal of quality, i.e., when a title enters a bestseller-list, the title is displayed more prominently on the retailers’ websites (e.g., a lower ranking in the search results).

Our findings have various economic as well as political implications. From the viewpoint of authors and publishers, our findings stress the importance of the brick-and-mortar channel. The odds of being discovered seem to be higher

in brick-and-mortar stores than in e-Commerce so that it is crucial that a book is presented in bookstores. From the viewpoint of (pure) online retailers, our findings show that there is still room for improvement when it comes to the promotion of unknown titles. This is of course an ambitious task because predicting the success of a new title, e.g., written by an unknown author, is extremely difficult. While originally applied to movies Goldman’s law, which says that “nobody knows anything”, also applies to books (Waldfogel, 2018, p.3-4). Especially machine-learning seems to be a promising way to enhance the predictability of the success of a title, although these approaches are (yet) at a relatively early stage (see, e.g., Wang et al. (2019)). From the viewpoint of policy makers, our results indicate that a high number of brick-and-mortar is an important driver of innovation in the book market. Given that books can be considered as merit goods (Appelman & Canoy, 2002), measures that increase the number of brick-and-mortar retailers can be socially beneficial. Such measures can, for instance, include fixed book price systems or subsidies.

Our paper is structured as follows. In Section 2, we compare the role of e-Commerce and brick-and-mortar retailers in the diffusion of new books. Section 3 provides an overview of our data set and contains some descriptive statistics. Our empirical analysis is presented in Section 4. In Section 4.1 we present a short theoretical exposition that provides the foundation of our identification strategy. Our estimation results are presented in Section 4.2. Section 5 concludes.

2 The Role of Brick-and-mortar Stores and E-Commerce in the Book Market

Consumer search for books can differ between the two retail channels e-Commerce and brick-and-mortar retailers for various reasons. Before investigating these, it seems appropriate to first stress some important aspects of books. Books can be seen as horizontally and vertically differentiated goods whose quality can only be asserted after consumption (i.e., after the consumer read the book). Such goods are sometimes referred to as experience goods (Nelson, 1974), which means that there is substantial uncertainty prior to purchase.

There are multiple measures that can be used to solve the problems arising from this uncertainty such as price-signalling (Bagwell & Riordan, 1991). As explained above, an external signal such as expert opinion is also an appropriate measure in the book market to reduce the consumers’ uncertainty (Clement et al., 2007). In general, expert opinion might affect demand in two ways. First, consumers may learn that a certain type of good exists (Hilger et al., 2011). Second, the opinion can act as a signal of quality (Reinstein & Snyder, 2005). A

reputable book dealer in a nearby book store may offer such expert opinion. Also the mere choice of which books to sell may act as a signal of quality if the book seller has a reputation to only stock books that satisfy a certain quality standard (Marvel & McCafferty, 1990).

An important measure to help consumers find their desired products in e-Commerce are recommendation systems. There is evidence that such measures increase sales Chen et al., 2004. However, there is evidence that there is limited attention to quality signals among consumers in e-Commerce (Watson et al., 2018). This could manifest in consumers' clicks being concentrated on the first few items of a search query. A similar behavior can be observed online search.¹ These systems also include customer reviews. The empirical evidence on how customer reviews affect sales is mixed (see, e.g., Watson et al. (2018)).

Another reason why the online and the offline channel may not be perfect substitutes is ad-hoc sales. While consumers may use online search engines to find the best prices for specific products (see, e.g., Tang et al. 2010), the number of impulse purchases in the offline channel may be higher (see, e.g., Burt and Sparks 2003). B&M retailers also have the opportunity to provide showrooms where consumers can physically investigate a product, which might positively influence demand (Bell et al., 2017). Although online retailers such as *Amazon.com* offer the opportunity to read the first few pages of a book for free, it may still be the case that some consumers may prefer to investigate the printed book.

3 Data and market insights

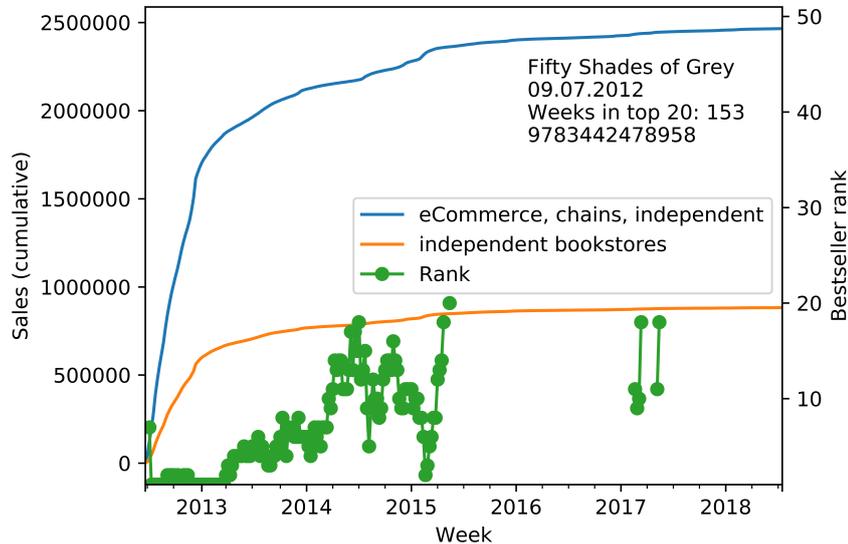
The German news magazine SPIEGEL publishes weekly top 20 bestselling titles by genre and format. This includes different lists for paperback, paperback plus² and hardcover editions as well as for fiction and non-fiction books. Besides the publicly accessible top 20 bestseller lists, SPIEGEL offers pay wall protected access to extended lists (top 50) for hardcover and paperback. We have sales data for 5,250 top 50 bestselling titles that have been published between 2011 and 2018. The data is provided by a German market research company called mediacontrol (MC) and is collected from POS systems (scanner data) of more than 9,000 outlets. We observe weekly sales and turnover by different sales channels. Among others, we are able to distinguish between sales in independent book stores, chain stores as well as sales of online retailers, such as Amazon.de. Figure 1 shows the

¹See, for instance, <https://www.searchenginewatch.com/2012/10/10/53-of-organic-search-clicks-go-to-first-link-study/>.

²Note that a paperback is defined different in Germany. The English word paperback refers to the German "Taschenbuch" (pocketbook), whereas the print and cover quality of the German "Paperback" can be denoted between hardcover and paperback. We will therefore refer to it as paperback plus.

cumulative sales of an exemplary book. In the following sections we will group sales as follows: (1) offline, i.e. chain stores and independent bookstores and (2) online, i.e. eCommerce.

Figure 1: Cumulative sales by sales channel



Based on consumer survey data of MC, we can tell that the turnover based market share of online retailers, such as Amazon.de, are higher for non-fiction than for fiction. In 2018, 18 percent of total fiction books turnover relates to online sales, 24 percent to chain stores and 45 percent to independent book stores. For non-fiction books the shares are 23 percent (online), 20 percent (chain stores) and 45 percent (independent book stores).³ In contrast to countries such as the USA or UK, eBook sales have not kicked off in Germany. According to survey data of GfK⁴, the market share is growing but ranges between 5 percent (turnover) and 8 percent (sales).

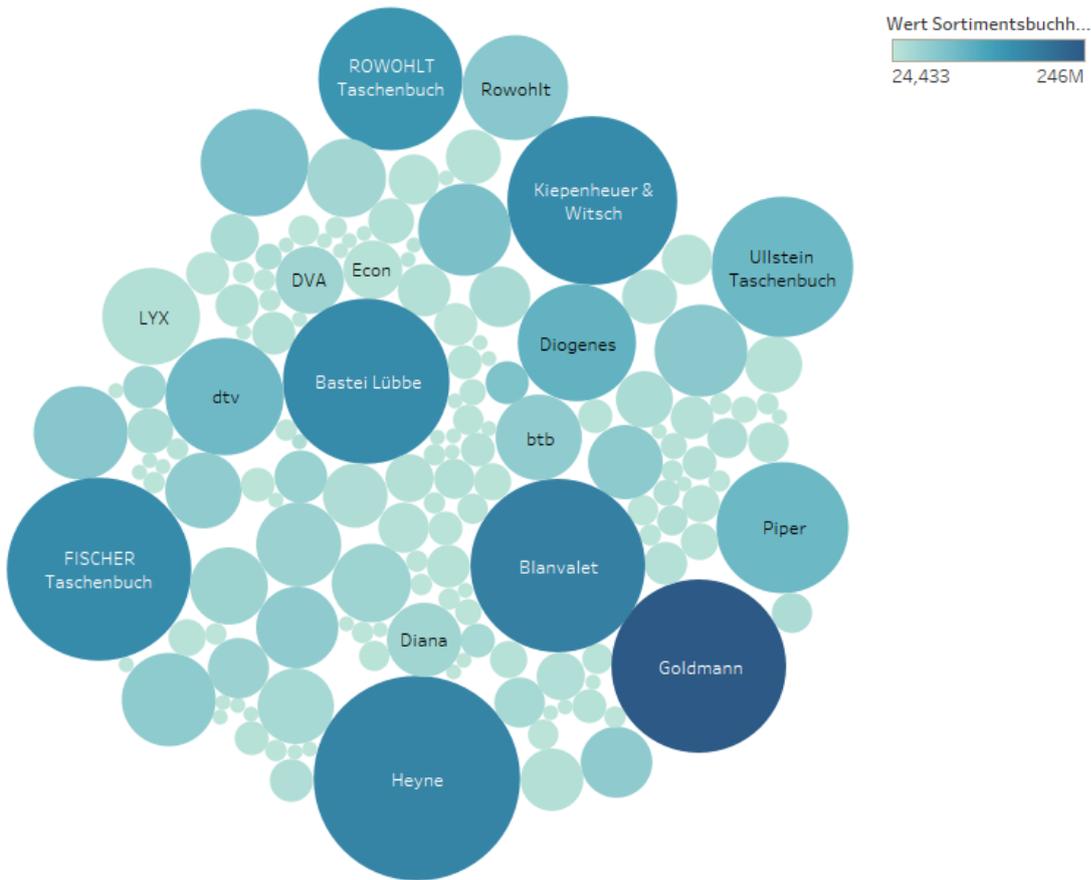
We can confirm the findings of Sorensen (2007) regarding the sales pattern of bestsellers for Germany, i.e. sales peak in the week of publishing and decrease over the following periods. For each sales channel, we observe the weekly top 50,000 best selling titles and are able to compare the SPIEGEL bestseller lists with the *true* sales ranking. Therefore, we can also confirm findings of Sorensen (2007) regarding the issue that bestseller lists are not solely based on sales ranks but other data and internal guidelines that are unpublished are used to prepare the lists.

Figure 2 shows the number of published Top 20 bestsellers (diameter) as well as the turnover (color) by imprint. Between 2011 and 2017 the imprint

³Due to data confidentiality agreements we are not able to provide market shares that are calculated from the scanner data.

⁴<https://de.statista.com/statistik/daten/studie/303339/umfrage/umsatzanteil-von-e-books-im-buchmarkt/>

Figure 2: Bestsellers by imprint



with the smallest turnover generated only 24,500 Euro from selling bestsellers, while *Goldman*, the publisher of the *Fifty Shades* trilogy in Germany achieved a turnover of roughly 246 million Euro. *Heyne*, part of *Random House*, published 180 bestsellers, while there are several imprints with only one Top 20 fiction or non-fiction bestseller between 2011 and 2017.

Table 1: Summary statistics

	N	Weeks Top 50	Copies	Turnover	Price
Direct Top 20	1,754	18.56	94,987.62	1,378,031.20	14.51
Late Top 20	849	20.70	77,180.44	1,116,297.60	14.46
Only Top 21 to 50	2,645	3.92	23,678.90	314,345.90	13.28

Table 1 shows some descriptive statistics on the bestsellers contained in our dataset. 1,754 books went directly to the top 20 bestseller list, while 849 books made it to the top 20 after first being ranked top 21 to 50. 2,645 books have never made it to the top 20 and, thus, serve as a control group. Retail prices of German books are set by publishers. Retailers are not allowed to deviate. The

retail price maintenance (RPM) benefits our analysis, i.e. we can treat prices as title individual fixed effects.

4 Empirical Analysis

4.1 Identification strategy

Our goal is to investigate whether and, if so, which, retail channel has an advantage in discovering commercially successful innovations in the book market, i.e., new bestseller-titles. In order to do so, we employ diffusion curves as established by Bass (1969). This model is used as a basis for our identification strategy because it was designed to analyze the adaption of innovation. It has also been applied to book sales before by, e.g., Beck (2007).

Following Mahajan et al. (2000, pp. 4–5), the basic idea of the model is as follows. Cumulative sales (i.e., the cumulative number of adopters) as a function of (continuous) time t are denoted $X(t)$. The change in the cumulative sales $x(t) := X'(t)$ can be understood as additional sales as a function of time t . These additional sales $x(t)$ can be described by the function

$$x(t) = \left(p + q \frac{X(t)}{m} \right) (m - X(t)). \quad (1)$$

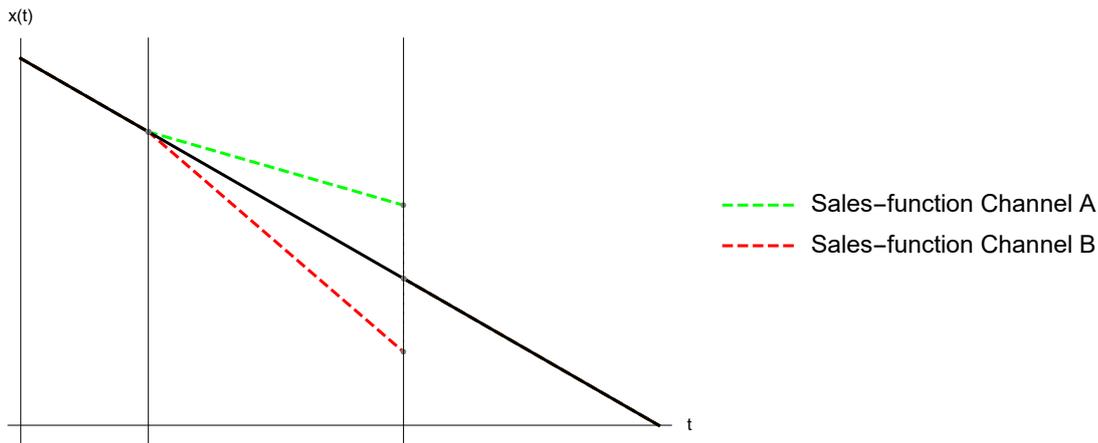
In equation (1), the parameter m captures the market potential. The parameter p is referred to as the coefficient of innovation or external influence and the parameter q as the parameter of imitation or internal influence. With respect to the book market, the parameters can be interpreted as follows. The higher the parameter p the higher the innate potential of a new title, which can be influenced by, e.g., the fame of the author or whether the title is a sequel. The parameter q measures the degree of “contagion”, i.e., how strongly current sales affect future sales, for instance due to word-of-mouth Beck (2007).

In most approaches, the parameters (m, p, q) are estimated based on existing data (see, e.g., Mahajan et al. (2000) for more information). For cases where $p \geq q$ sales peak in the period of release, and monotonically fall in all subsequent periods. Anticipating our results, our average title shows such a pattern. We therefore focus on that case throughout the following explanations.

We use the basic idea of the Bass-model to identify which retailer benefits the most from the exogenous quality signal of a title entering a bestseller-list. The idea is illustrated in Figure 3.

In Figure 3, the solid curve depicts additional sales x as a function of time t for a title that never becomes a bestseller. As explained above, sales x peak in the release period and fall in all subsequent periods. To better illustrate our results,

Figure 3: Sales figures in the Bass-model



$x(t)$ is presented as a linear function, although sales curves in the Bass-model are non-linear functions of time. The vertical lines represent the time span in which a title appears on the bestseller list. Let t^* denote this interval of periods where the title is in the bestseller-list.⁵

The dotted, green line in Figure 3 depicts sales of channel A. This channel benefits from the quality signal, i.e., that a title enters the bestseller list, in a sense that sales decrease more slowly compared to a title that is not on a bestseller-list. In contrast, in comparison to a title that is not on the bestseller-list sales in channel B decrease faster. This implies that consumers in channel B are *ceteris paribus* more likely to read titles which are not on bestseller-lists. Consequently, consumers in channel B either (i) purchase titles future (but not current) bestsellers or (ii) read books which never become bestsellers. The crucial question we want to answer is whether the online or the offline channel shows the properties of channel A or channel B.

To formalize this idea, define by $x_{\text{off}}(t)$ sales in brick-and-mortar stores and by $x_{\text{on}}(t)$ sales in e-Commerce. The change in additional sales after a title enters a bestseller-list is then $\frac{dx_i(t^*)}{dt}$, for $i \in \{\text{off}, \text{on}\}$. If, for the average title, it holds that $\frac{dx_{\text{off}}(t^*)}{dt} < 0$ and $\frac{dx_{\text{on}}(t^*)}{dt} > 0$, compared to a title that is not on the bestseller-list, offline sales decrease more strongly in periods where a title is on the bestseller list whereas online sales decrease less strongly in those periods. In that case, as explained above, one could argue that consumers who purchase in brick-and-mortar stores only purchase titles which are not yet bestsellers, old bestsellers or titles which never become bestsellers. If the signs are reversed, the interpretation changes and consumers in the online channels are those who purchase unknown titles which are not (yet) on bestseller-lists.

⁵Formally, let t_0 and t_1 be the points in time when a title enters and exits the bestseller-list, respectively. Then, $t^* := \{t : t \in [t_0, t_1]\}$.

4.2 Estimation

We are using an approach introduced by Sorensen (2007), who basically assumes that sales follow an AR(1)-Process:

$$sales_{it} = \lambda_{it}sales_{it-1} + \epsilon_{it} \quad (2)$$

$$\lambda_{it} = X'_{it}\lambda \quad (3)$$

The set of covariates X_{it} includes dummy variables that indicate whether a book has been listed on one of the top 20 bestseller lists (fiction or non-fiction) in week t and whether a review has been published in one of the six major newspapers in this week. t corresponds to weeks since a title has been published, ranging from $t = 0$ to $t = 78$.⁶ We furthermore include a time trend to account for declining sales pattern. To control for seasonal and book specific effects, we include week and year fixed effects as well as book fixed effects.

Table 2: Results

	(1) Both channels	(2) Offline	(3) Online
Bestseller	-0.107*** (-3.96)	-0.188*** (-4.67)	0.126** (2.82)
Review	-0.0385 (-0.80)	-0.0509 (-0.93)	0.0185 (0.40)
Weeks since release	-0.00650*** (-4.03)	-0.00900*** (-5.65)	0.00344 (1.03)
Weeks since release squared	-0.0000415** (-2.82)	-0.0000146 (-0.97)	-0.000125*** (-6.99)
N	370,193	370,193	370,193
R^2	0.957	0.953	0.918

t statistics in parentheses. All explanatory variables are interacted with lagged sales.

Book level clustered standard errors are shown in parentheses.

Week, year and book fixed effects included.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

In Table 2, the first row depicts the coefficient for the dummy-variable that becomes one when a title is on the bestseller-lists. One can see that that this decreases the sales patterns for an average title throughout all channels by approximately 11%. This effects becomes even stronger (-19%) when only the offline channel is taken into account. However, in the online channel the sales pattern increases by approximately 13%.

⁶The RPM is valid for 18 months, this corresponds to about 78 weeks. After that prices may be adjusted.

To get a better understanding of this result, it is useful to investigate the coefficients for *Weeks since release* and *Weeks since release squared*. One can see either one or both coefficients are significantly negative, which means that as soon as the average title is released sales start to decline. Against this background, the above results imply that sales in e-Commerce decline approximately 13% more slowly whereas offline sales decline approximately 19% more strongly compared to a title that is not on the bestseller-lists. The latter result is in contrast to Sorensen (2007), who only finds a positive coefficient over all sales when a title is on a bestseller-list.

5 Conclusion

This paper studies the roles the online and offline retail channels play in the discovery of new bestsellers. In doing so, we have analyzed how sales in e-Commerce and brick-and-mortar stores react when titles become bestsellers, which can be understood as a signal of that title's quality. We find that when the signal occurs for an average title online sales increase and offline sales decrease compared to a title that is currently not on the bestseller-lists. In particular, because we model sales as an auto-regressive process, sales decline less strongly in e-Commerce and decrease more rapidly at brick-and-mortar retailers compared to titles that are not on the bestseller-lists.

These findings imply that online sales respond positively to the quality signal of a title entering the bestseller-lists. The offline channel shows the opposite effect, which means that in the offline channel a title is already widely known when it enters the bestseller-lists. This could be interpreted as an advantage of the offline channel in the discovery of new bestsellers. *Ceteris paribus*, consumers may purchase more non-bestselling titles in the offline channel. Although among these non-bestselling titles there might be some which never become bestsellers or which have already been on older bestseller lists, some of these will become *future* bestsellers. Given that the market share of offline retailers is about three times larger than the market share of online retailers, our findings could indicate that brick-and-mortar stores foster innovation in the book market.

These findings can be explained by demand-side or supply-side effects. On the demand-side, the majority of heavy book readers purchases offline. Given that they read a particularly large number of books the odds that they discover new bestsellers are relatively high. It could also be the case that consumers in e-Commerce do not become aware of new and commercially promising titles which are not (yet) bestsellers. It could, for instance, be the case that most consumers purchase online when they already know what they want to buy whereas they browse for new products offline. Related to that, on the supply-side of the market,

it could be that online retailers choose to promote titles more excessively when the quality of a title is already clear, i.e., when the title is already a bestseller. This, again, could lead to more consumers purchasing bestseller-titles online because of the more prominent product placements on the online retailers' websites.

Our results also indicate that there exists a positive externality that arises from brick-and-mortar retailers and which online retailers benefit from. If a title enters the bestseller-lists primarily due to offline sales, online retailers benefit from that because the uncertainty about a title's profitability is reduced. As a consequence, online advertising can be used more efficiently, which is beneficial from those the perspective of e-Commerce as well as consumers who predominantly purchase online. Although we observe that this might happen on average, there are certain cases in our data where titles are predominantly sold in e-Commerce before they become bestsellers. From a broader perspective, one could hypothesize that the two channels are complementary in a sense that more new bestsellers are discovered relative to a case where only one channel exists. This has to be investigated in future research.

Additionally, from a policy perspective, if governments want to increase the innovativeness in the book market, they can choose measures which increase the number of brick-and-mortar retailers. This can, for instance, be through direct subsidies or by installing fixed book price systems, which tend to particularly benefit smaller retailers (see, e.g., Fishwick et al. (1997)). These are most likely only active in the offline-market.

References

- Aguiar, L., & Waldfogel, J. (2018). *Platforms, promotion, and product discovery: Evidence from spotify playlists* (tech. rep.). National Bureau of Economic Research.
- Appelman, M., & Canoy, M. (2002). Notes and communications, horses for courses: Why europe should not harmonise its book policies. *De Economist*, *150*(5), 583–600.
- Bagwell, K., & Riordan, M. H. (1991). High and declining prices signal product quality. *The American Economic Review*, 224–239.
- Banerjee, A. V. (1992). A simple model of herd behavior. *The quarterly journal of economics*, *107*(3), 797–817.
- Bass, F. M. (1969). A new product growth for model consumer durables. *Management science*, *15*(5), 215–227.
- Beck, J. (2007). The sales effect of word of mouth: A model for creative goods and estimates for novels. *Journal of Cultural Economics*, *31*(1), 5–23.
- Beck, J. (2012). Advance contracting, word-of-mouth, and new-product success in creative industries: A quantification for books. *Journal of Media Economics*, *25*(2), 75–97.
- Bell, D. R., Gallino, S., & Moreno, A. (2017). Offline showrooms in omnichannel retail: Demand and operational benefits. *Management Science*, *64*(4), 1629–1651.
- Burt, S., & Sparks, L. (2003). E-commerce and the retail process: A review. *Journal of Retailing and Consumer Services*, *10*(5), 275–286.
- Chen, P.-Y., Wu, S.-y., & Yoon, J. (2004). The impact of online recommendations and consumer feedback on sales. *ICIS 2004 Proceedings*, 58.
- Clement, M., Proppe, D., & Rott, A. (2007). Do critics make bestsellers? opinion leaders and the success of books. *Journal of Media Economics*, *20*(2), 77–105.
- Fishwick, F., Fitzsimons, S., & Muir, L. H. (1997). *Effects of the Abandonment of the Net Book Agreement: First Interim Research Report*. Book Trust London.
- Hilger, J., Rafert, G., & Villas-Boas, S. (2011). Expert opinion and the demand for experience goods: An experimental approach in the retail wine market. *Review of Economics and Statistics*, *93*(4), 1289–1296.
- King, T. (2007). Does film criticism affect box office earnings? evidence from movies released in the us in 2003. *Journal of Cultural Economics*, *31*(3), 171–186.
- Mahajan, V., Muller, E., & Wind, Y. (2000). *New-product diffusion models* (Vol. 11). Springer Science & Business Media.

- Marvel, H. P., & McCafferty, S. (1990). Resale price maintenance and quality certification. *J. Reprints Antitrust L. & Econ.*, 20, 251.
- Nelson, P. (1974). Advertising as information. *Journal of Political Economy*, 82(4), 729–754.
- Reinstein, D. A., & Snyder, C. M. (2005). The influence of expert reviews on consumer demand for experience goods: A case study of movie critics. *The journal of industrial economics*, 53(1), 27–51.
- Sorensen, A. T. (2007). Bestseller lists and product variety. *The journal of industrial economics*, 55(4), 715–738.
- Tang, Z., Smith, M. D., & Montgomery, A. (2010). The impact of shopbot use on prices and price dispersion: Evidence from online book retailing. *International Journal of Industrial Organization*, 28(6), 579–590.
- Waldfoegel, J. (2015). Digitization and the quality of new media products: The case of music. In *Economic analysis of the digital economy* (pp. 407–442). University of Chicago Press.
- Waldfoegel, J. (2016). Cinematic explosion: New products, unpredictability and realized quality in the digital era. *The Journal of Industrial Economics*, 64(4), 755–772.
- Waldfoegel, J. (2018). *Digital renaissance: What data and economics tell us about the future of popular culture*. Princeton University Press.
- Wang, X., Yucesoy, B., Varol, O., Eliassi-Rad, T., & Barabási, A.-L. (2019). Success in books: Predicting book sales before publication. *EPJ Data Science*, 8(1), 31.
- Watson, J., Ghosh, A. P., & Trusov, M. (2018). Swayed by the numbers: The consequences of displaying product review attributes. *Journal of Marketing*, 82(6), 109–131.