The effect of shocks in leisure time on book consumption: evidence from Goodreads

Javier Torralba Flores

June 20, 2023

ANR: 146430

SNR: 2042878

Date of submission: June 20, 2023

BSc thesis for Economics

Name of supervisor: dr. Lachlan Deer

Second reader: dr. Elena Cettolin

Number of words: 7,965

Contents

1. Introduction

People use their leisure time to enjoy different activities, such as watching movies, spending time with friends, or reading books. Increased leisure time for people means more time for consumers to use leisure goods. However, it is not clear how consumers allocate this extra time to leisure goods. There may be an exploration mechanism, where consumers do not consume the leisure good they usually do anymore. At the same time, there may be a preference discovery mechanism, where people consume the same type of leisure good, but with slight variations.

For at least two years, the COVID-19 pandemic reshaped the way people live their lives. While the pandemic changed several of society's usual ways of functioning, a key one was how people used their time. While in everyday society, people had the chance to socialize in person with friends and family, commute to work, and spend time at cinemas, restaurants, or at the gym, all these aspects of life, and many more, changed almost overnight. COVID-19 restrictions imposed by governments served as a shock to how people use their free time. Restrictions led to increased leisure activities alone (Lee & Tipoe (2021)).

My research uses the COVID lockdowns as a shock that increases leisure time to study questions using books as a leisure good to study. More specifically, the problem statement of my study is *"The academic literature related to reading behavior during the COVID pandemic is scarce, with unclear effects of leisure time shocks on book consumption".* I will investigate the research question, "*How do consumers leverage extra leisure time to book-related consumption on Goodreads?".* At the same time, I look to answer the following sub-questions *"Do consumers of the platform Goodreads predominantly adopt an exploration or preference discovery mechanism when there is a positive shock to their leisure time?"* and "*Does extra leisure time make consumers more critical of the products they consume?"*

Publishing houses, authors, bookstores, online platforms, and many more are all parts of the customer's journey when consuming books. It is estimated that book sales revenue for 2023 alone is $78.07 billion, according to global book sale statistics (Curcic, 2023). Goodreads is one of the world's largest sites for readers and book recommendations (Goodreads, 2023). The site alone counts 125 million users (Michelle (2022)). Given the magnitude of the website and of the book industry, understanding how COVID restrictions and other potential leisure time shocks affect book consumption could be beneficial for several parties.

I obtained data from the platform Goodreads and combined this with the Oxford COVID-19 Government Response Tracker (OxCGRT) (Hale et al. (2021)). Using this data, I created an econometric model in the form of an OLS regression with fixed effects to evaluate the effects of leisure time shocks on the consumption of books. With this, I was able to find that users increase their book consumption in the form of reading more pages, books, and doing this quicker as restrictions increase. At the same time, users read older books and extra leisure time does not affect the way users rate books.

The remainder of the thesis proceeds as follows. In chapter 2, before delving into the literature review, the reader is contextualized on the background of the book industry, COVID and Goodreads. Chapter 3 reviews the relevant literature. Chapter 4 describes the data used. Chapter 5 gives model-free evidence that visualizes the data. Chapter 6 shows the set-up of the model used to apply the inferential analysis. Chapter 7 shows the results. Chapter 8 discusses the implications of the findings and concludes the paper.

2. Background: book industry and COVID-19

This chapter reviews background and impact of the pandemic on the book industry and explains how the Goodreads site works. This section is presented before the literature review section to allow the reader to familiarize with book industry and the social media platform, Goodreads, as it is worthwhile to be contextualized with this before delving into the literature review.

2.1 The book industry and the COVID-19 Pandemic

Several news outlets and popular magazines reported soaring sales by booksellers during months of the pandemic. For instance, an article from the BBC published in October of 2021 (Bloom (2021)) reported that sales of physical books rose strongly during 2020 and 2021. They report how many brick-and-mortar stores adapted their business models to "click and collect" services, which helped them fight a potential downturn in sales due to COVID. Similarly, Printweek magazine reported that the UK book market grew 5.2% by volume in 2020 (Stuart-Turner (2021)). Private non-profit organizations such as the World Economic Forum also reported an increase in book sales at the beginning of the pandemic (Charlton (2020)).

A report from the Federation of European Publishers in 2021 gives insight into the situation (Federation of European Publishers (2021)). While the report starts by explaining the fears of a dire economic situation for publishers, expecting a loss in the sector of 15-25%, the loss in Europe appears closer to 2-5%. There are several factors for this, such as strong sales over the summer months of 2020, limited leisure activities due to travel restrictions, and reader solidarity towards book sellers. Some countries saw a surge in publishing sales. Retailers saw a shift from brick-and-mortar stores to online retailers, with sellers in countries with better online sales infrastructure being benefitted most.

2.2 Goodreads

Goodreads is one of the world's largest sites for readers and book recommendations (Goodreads (2023)). The platforms can be described as a social media platform for books, where users can share the books they have read or plan to read and share this with their friends. Goodreads was launched in January 2007 and after reaching 16 million users in 2013, Amazon bought the platform (Flood (2013)). Today, it has over 125 million members with a database of 3.5 billion books (Michelle (2022)). Figure 1 shows the number of users on Goodreads over time from 2011 to 2019.

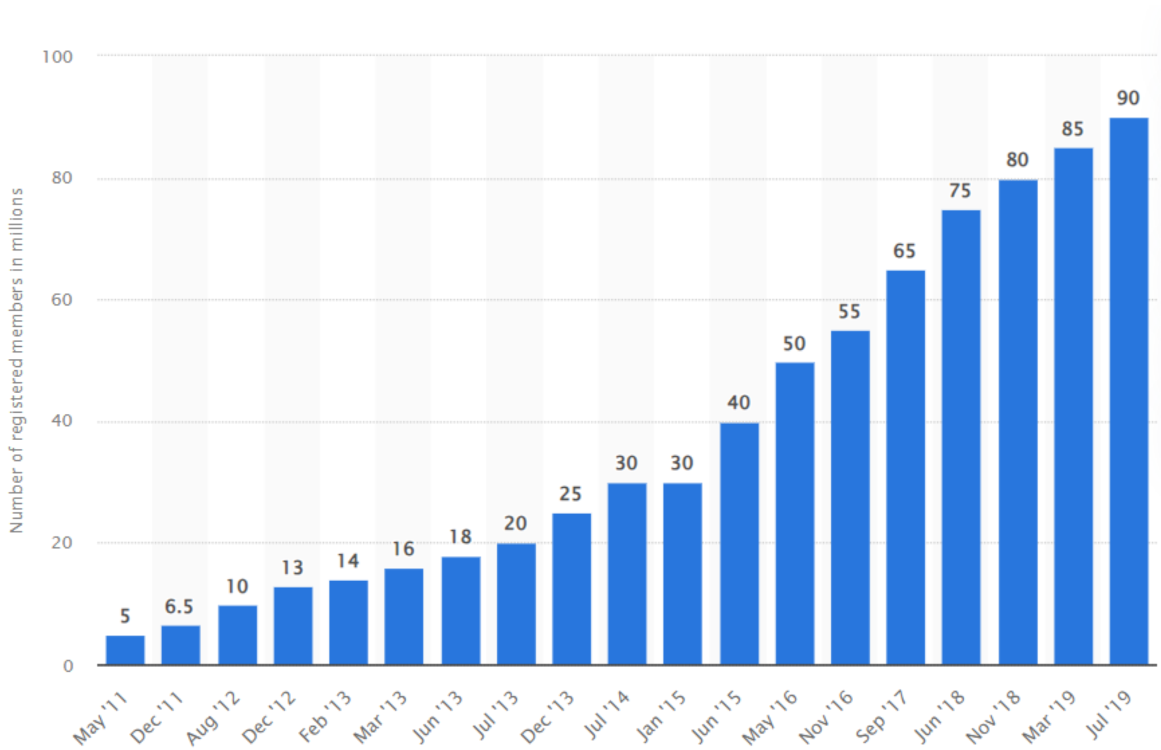


Figure 1: number of users (in millions) in Goodreads (Statista (2023))

Goodreads users self-report the books they wish to read along with the dates they started and finished reading a book. Users can give ratings and reviews to books. It is also possible to join the site's discussion boards, groups, and other social media-like features, such as making friends and sharing book recommendations to them. The home page of the platform is a feed such as what you would find in platforms like Facebook or Instagram, as figure 2 shows.

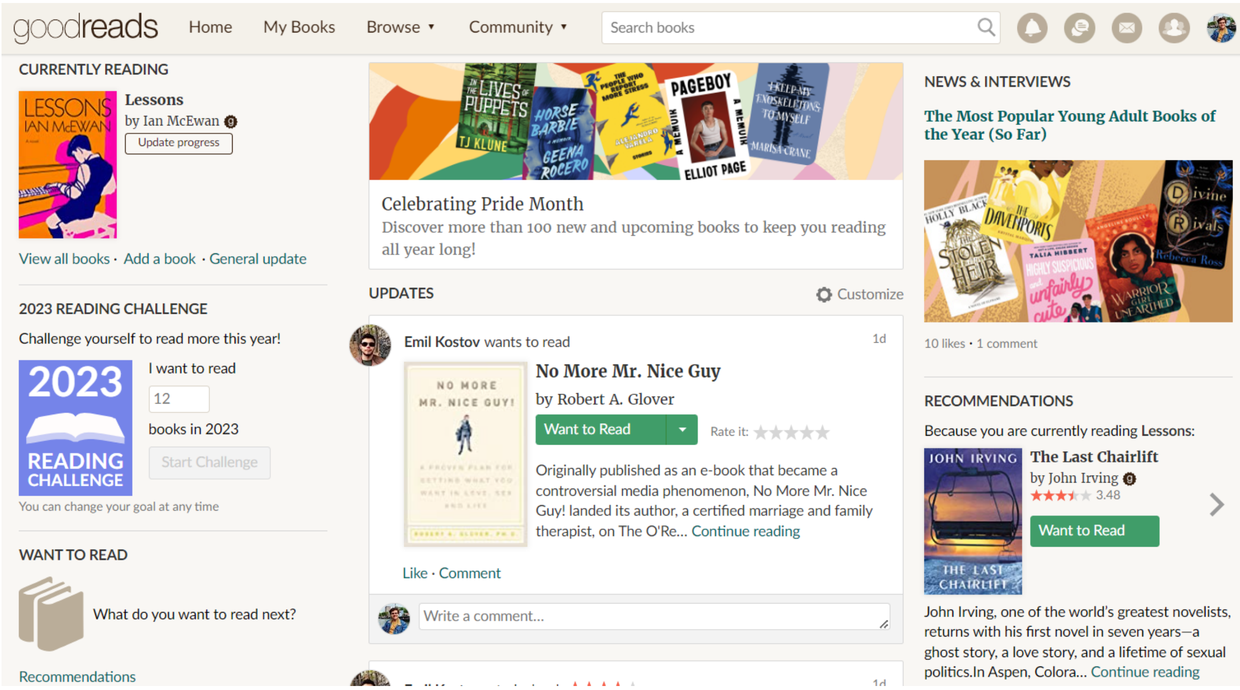


Figure 2: Main feed Goodreads

Knowing the size of the platform with more than 100 million users, along with its capabilities to understand the user's social and reading behavior, Goodreads can be a great source of data for different fields of research. Specifically, this data can be useful for authors, publishers, and retailers, who could all be benefited from understanding social and reading patterns of readers who consume their products.

3. Literature review

This chapter is divided into four sections: first a review of literature related to consumption during the COVID-19 pandemic, as it is important to understand how consumption of goods may be affected as a whole due to the pandemic. Second, a review of literature related to book consumption and the book industry, as it allows us to understand how the industry works and how it was affected by the pandemic. Third, a review of literature related to the platform Goodreads, as it shows how other research has used Goodreads data and influences the way I will use this data to understand the effects of shocks to leisure time on book consumption. Fourth, a hypothesis formulation that uses the literature reviewed as arguments to this formulations.

3.1 Literature Review on overall consumption patterns and COVID

It is important to understand the literature on how consumers adapted change their habits helps shed light on the societal and economic context. I put together a summary of findings from papers related to this literature in this chapter. Many of these papers focus on consumer spending during COVID and show that overall spending and consumption were a negative affected. Namely, Chen et al. (2020) show how overall offline consumption dropped by an average of 32% in China during the three-month post-outbreak period. Similarly, Chetty et al. (2020) show a 74.7 % reduction in personal consumption expenditures in the second quarter of 2020, during the pandemic's beginning.

Oblander & Mccarthy, (2023) studied the effects of COVID-19 on the short- and long-term effects of consumption patterns due to COVID-19. They estimate the impact of COVID-19 on purchasing behavior in 12 consumption categories. These categories involved both online and offline businesses. In short, all 12 categories were greatly affected. They found significant shifts in consumer behavior early in the pandemic, namely a strong substitution effect, with online categories seeing a massive boost in sales, with the opposite effect for offline categories. At the beginning of the crisis, online categories grew the most, with impulsive and irregular purchases followed by a tapering off. At the same time, more routine, subscription-based purchases did not enjoy an immediate boost but a slow, gradual increase over time. Their findings show that some consumption categories shifted towards online shopping rather than decreased overall consumption in all potential categories.

Research of the effects of COVID restrictions on the music industry closely align with this thesis (Sim et al. (2022)). It is relevant to understand the type of leisure time shock that COVID restrictions represent, mainly restrictions on mobility and stay-at-home orders. The authors demonstrate how individuals reduced their music consumption, mainly due to a substitution effect. Users no longer need to commute; they reduce their levels of exercise and mobility, all of which contribute to a reduction of music consumption and incite a substitution towards visual forms of entertainment, such as video streaming. Like the data I present, the authors used industry data (data from streaming the top 200 songs on Spotify) and the Oxford COVID-19 Government Response Tracker (Hale et al. (2021)). They perform a differences-in-differences analysis to understand the effects of restrictions on music consumption. However, the authors do not delve into how different levels of restrictions affect music consumption differently.

Even though plenty of academic literature touches on the effects of COVID restrictions on consumer spending and consumption, little to no literature focuses on book consumption patterns. To the best of my knowledge, no literature explains how different levels of COVID restrictions, which represent an increase in leisure time, affect the consumption of books. This is important to understand because it sheds a light on how people may consume leisure goods differently when their leisure time increases because of COVID restrictions. This paper aims to fill that literature void.

3.2 Literature Review on book consumption and book industry

Given that I will study the effects of leisure time shocks on book consumption, it is important to understand the relevant literature on the state book industry and book consumption. The literature I present focuses on three main topics. Firstly, literature covering the characteristics of more popular books and how a digital social environment might influence the popularity of books (Ding & Li (2019)). Secondly, literature that aims to understand the consumer behavior on the e-book market (Liao & Liu (2022)). Thirdly, the way the COVID pandemic affected book consumption and the book industry (Guren et al. (2021)).

Ding and Li (2019) study the herding in the consumption and purchase of digital goods. They study whether users engage in herding behavior and identify factors that might influence this. While they study several goods, one of them is books. The authors find that books that appear at the top of lists on websites that sell books, being categorized as fiction and with highly reputable authors influence book's clicks and sales. The authors also find that users engage in herding of consumption of digital goods such as online sales of books. This is relevant to my research it might describe the behavior of book consumers.

Liao and Liu (2022) have studied the effects of different marketing strategies on e-book consumption. They highlight the fact that widespread adoption of mobile devices has significantly increased the sales of e-books and decline of print books in North America. Many retailers in the U.S have adopted a business model of dual publication modes, such as Amazon (the owner of Goodreads) selling both physical books and e-books. This is related to my research as Goodreads is a social media platform that might influence book buyers' decisions. The authors found that several factors, such as giving free e-book previews, including a numbered time reference in the e-book title, and consumer's expertise of technology influence consumers' buying decisions of e-books.

Guren et.al (2021) study the impact of the COVID pandemic on book publishing. They find that even though the pandemic posed as a threat to the book publishing industry, publishers fared well in 2020, escaping many of the pandemic related disruptions faced by several other sectors. Most publishers experienced revenues in 2020, but this were mostly influenced by the short-term social context rather than long-term investment opportunities. Because of public restrictions on education, the demand for digital educational content increased, leading to a good year for educational book publishers. Finally, given the retail landscape and its shifts due to the pandemic, publishers are now advised to prioritize online and digital sales over physical bookstores, linking well to the findings in chapter 3.1.

All in all, this literature gives us insight to the way people consume books, marketing strategies that might influence book retailers and the impact of the pandemic on book publishers and consumers. This is important to know as it relates directly to the leisure good I focus on my research, which is books.

3.3 Literature Review on Goodreads

As I use Goodreads data in my research, it is important to understand the literature on how other research has used similar Goodreads data. In addition, understanding the way other papers have used and studied this data, helps shape the model I present in chapter 6. The literature I review ranges from understanding how the Goodreads platform is used by its users (Thelwall & Kousha (2017)) to understanding how individuals can learn from social ties (Zhang & Godes (2018)) or the welfare effects of gender-inclusive intellectual property creation (Waldfogel & Kappel Chair (2023)). Ahead I describe, compare, and contrast this relevant literature to my research.

Thelwall & Kousha (2017) study how Goodreads users interact with the platform. Their study focused on the behavior of 50,000 members and aimed to provide a descriptive analysis of how users interact with Goodreads. Their research revealed various characteristics of users on Goodreads. For example, they found that older users are more active. Additionally, they discovered no clear relationship between social activity on the platform and book consumption reporting, which includes how people review books. Lastly, the authors found little evidence of gender differences in platform usage. Like Thelwall and Kousha's research, my paper also aims to understand user behavior on Goodreads, focusing on characterizing different types of users on the platform.

Research using Goodreads data goes beyond understanding the platform. Zhang & Godes (2018) use this type of data to study the impact of social ties on learning from close and distant social relationships. Their research presents results on how more social ties can lead to better decisions. The authors find that consumers with more friends on the platform post higher ratings than those with fewer friends. This is relevant to my thesis as different types of users may be critical towards rating books, which will be important to control for in the model presented in chapter 6. More literature on Goodreads data is presented by Waldfogel & Kappel Chair (2023). They looked to determine the welfare effects of gender-inclusive intellectual property creation, documenting the growth in female authorship and how this has helped consumers. The data they use contains about 230 million interactions with 2.3 million books by 800,000 users on Goodreads between 2007 and 2016.

All in all, the relevant research related to Goodreads will help me construct a part of the model I present ahead. Findings from Zhang & Godes (2018) and Thelwall & Kousha (2017) relate to how users that have more friends on the platform give higher ratings to books and how older users are more active than newer users on how different characteristics of users may be correlated with how members use the platform will help shape fixed effects I present in the model ahead. At the same time, to the best of my knowledge, no published literature studies the effects of a shock on leisure time and its various levels of disruption on book consumption. This thesis aims to fill that void.

3.4 Hypothesis formulation

Given the literature on COVID, book consumption and industry, and Goodreads, I can come up with three hypotheses related to how leisure time shocks affect book consumption. First, I hypothesize that *book consumption will increase.* Based on the research of Sim et al. (2022), they find that music consumption is reduced due to people traveling less and living a more sedentary life. While this might be the case for a leisure good like music, I hypothesize that we will see the opposite effect on books, as they are a more sedentary activity that does not require socializing. Also, as it has been documented by several papers (Liao and Liu (2022), Oblander & Mccarthy, (2023), Guren et.al (2021)), the use of online platforms for consumption and the use of e-books increased, as well as parts of the book industry being benefitted. This might also hint to an increase of book consumption reflected on the platform Goodreads. Second, I hypothesize that given that people read more books, *users will read older books*. As there is a limited supply of books in the market, and users read more books, they will start reading older books. My third hypothesis is that *users will not change the way they rate books as their leisure time increases*. As

Tehelwall and Kousha show in their paper, there is a relatively low correlation between the number of books read and the number of reviews written on Goodreads, which hints towards users not rating books differently when there is extra book consumption.

4. Data

This chapter describes how I obtained, altered, and used the data to conduct my research. I use data from two primary sources, Goodreads, and the Oxford COVID-19 Government Response Tracker (OxCGRT) Stringency Index. In the end, one can find an explanation and description of the relevant dependent and independent variables and the specifications of the final dataset.

4.1 Goodreads data

I used data from the website Goodreads that I obtained from assistant Professor Lachlan Deer while working for him as a research assistant. The data contains information at both the user and book level. The user level data consists of general information about users in the platform, with 76,924 unique users. General information of the users consists of country the user lives in, number of ratings, number of reviews, number of books read, number of friends, age, gender, date joined, days active, the average number of books read per day, and number of books left to read that a user has. Users come from 31 countries in the data, which can be found in appendix A1. The books level data contains information about the books these users read. The book variables consist of the name of the book, date the user started and finished reading it, as well as when it was added to their library, the rating the user gave it, the number of pages it has, the average rating on a scale from 1 to 5, date the book was published, and the speed in days it took the user to finish reading the book.

4.2 Stringency index

The Oxford COVID-19 Government Response Tracker (OxCGRT) collects systematic information on government policy measured during the COVID-19 pandemic (Hale et al. (2021)). This project has tracked responses since 1 January 2020 by over 180 countries. I focus on the stringency index, which measures the strictness of lockdown-style policies that restrict people's behavior.

The stringency index is built with recordings of nine categories measuring the strictness of a lockdown in each country in each day. The final index is calculated by creating a weighted average of the ratio of each of the nine categories, which is also geographically adjusted. For instance, a policy targeted to the whole country will raise the stringency index more than a policy targeted at a single city of the country. The index also takes the conservative assumption that if there is no relevant data available for one or more of the nine ratios, that given ratio drops to a value of zero, which drags down the index. The stringency index ends up being a continuous variable ranging from 0 to 100, representing the strictness of the lockdown level for each day of the year for a country (Hale et al. (2021)).

This index was collected for all countries in the Goodreads data and merged by country and date read, such that the stringency index will match the date a person finished reading the book.

4.3 Adaptations to the Stringency Index

The COVID stringency index is of an ordinal scale, which means that the intervals between the values may not be equal or meaningful. For instance, an increase from 20 to 40 might not be of the same magnitude as an increase from 40 to 60. It might be the case that COVID restriction increasing from 20 to 40 represent a relatively small change in how society functions, whereas a change from 40 to 60 has a drastic change. This might be an issue when interpreting regression coefficients in the form of an OLS Fixed Effects regression. To tackle this issue, I disentangled and calculated the stringency index such that they would represent four levels of restrictions. Appendix A2 includes the calculations taken to construct the four categories of levels of restrictions. Here is an explanation of each of the intervals:

- [0] = <u>No restrictions</u>: Schools and businesses are operating as usual, with no restrictions on public events, transportation is running normally, and no recommendations to stay home or limit travel.
- [0.01 – 41.67] = <u>Light restrictions:</u> This includes recommendations to close or alter schools and businesses to operate differently. There are also recommendations to cancel public events and limit large gatherings. Transportation may also be affected, with reduced volume available. People may be urged to stay at home and avoid travel and quarantine measures may be in place for arrivals from certain regions. Public officials may emphasize caution about the COVID situation.
- [41.68 – 83.33] = <u>Closing of some of society:</u> This means that some schools or businesses may be required to close or alter their operations, and there may be restrictions on gatherings of up to 11-100 people. Internal movement restrictions may also be in place, and there may be a coordinated public information campaign to raise awareness of COVID. Border closures may also be in effect.
- [83.34 – 100.00] = <u>Closing of all but essential of society:</u> This represents the most stringent level of restrictions. It includes requirements to close all levels of education and non-essential workplaces, cancel public events, and restrict gatherings to no more than ten people. There may be prohibitions on leaving the house, except for limited exceptions, and internal movement restrictions may be in place. Border closures may also be in effect, and there may be a coordinated public information campaign to raise awareness of COVID across multiple media channels.

4.4 Final data set, dependent, and independent variables

4.4.1 Dependent and independent variables:

Dependent variables:

Five dependent variables will help evaluate book consumption the research questions and three hypotheses I proposed in chapters 1 and 3.4, respectively. Three variables evaluate the results at the book level, and two others at the user level. Ahead you can find an explanation for each of them.

Book-user level variables:

- Reading time per book: measures the users' reading speed on the platform. Users can self-report the date they started and finished reading a book. This variable is calculated by subtracting both dates and obtaining a numeric value representing the number of days it took the user to read the book. For example, if a user started reading a book on 2020-05-17 and finished reading it on 2020-05-18, this variable takes a value of 1.
- Age of book: measures the age of a book in years. The variable is calculated by calculating the difference between the year the book was read and the year it was published. For instance, if a book is read in 2018, that was published in 2016, the variable takes a value of 2.
- User rating: this variable refers to a user's book rating. Users can rate a book on a discrete scale from 1 to 5.

User-level variables:

- Number of books read by a user: measures the number of books a user reads in two weeks. As can be seen in Table 1, the average reading time per book is 11 days. Two weeks is an appropriate window of time to allow users to have enough time to read a book (or more). Later on in chapter 6, the measure of two weeks will also allow to account for seasonality.

- Number of pages read by a user: measures the number of pages a given user reads in two weeks. The same rationale for choosing to weeks as the number of books is applied to this variable.

Independent variable

The independent variable is the COVID OxCGRT stringency index divided into four intervals of restrictions. The four categories of the independent variable are, “No restrictions”, “Light restrictions”, “Closing of some of society”, “Closing of all but essential of society.”

4.4.3 Final dataset

The final dataset consists of 6,544,892 observations with 39 variables, starting from January 1, 2015, to December 31, 2021. Table 1 summarizes the descriptive statistics for the dependent and independent variables. There are several the variables mentioned in table 1 do not reach the number of observations of the full dataset. For reading time and user rating, users self-report this information, and because some users don’t give a rating, or don’t state the date they started reading a book, it is not possible to compute these variables, leaving some NAs in the data. For “Age of book” and “Number of pages read”, some books do not have data on the year the book was published, or the number of pages of the book. Finally, for “Number of total pages read” and “Number of total books read”, these variables are aggregated for two weeks for all users, therefore reducing the number of observations for these variables. Each row of data consists of a user self-reporting the date they finished reading a book, with the information described above about the user and book.

Table 1: Summary statistics for dependent and independent variables

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Median	Pctl(75)	Max
Reading time per book	5,445,881	11.189	23.685	0	1	3	10	206
Age of book	5,801,837	16.037	26.255	0	2	5	17	154
User rating	6,070,249	3.772	0.973	1	3	4	5	5
Number of total pages read	2,473,049	618.642	448.748	1	305	473	810	2,447
Number of total books read	2,684,989	2.163	1.676	1	1	2	3	10
Stringency Index	2,321,856	57.454	22.349	0.000	43.980	62.960	73.150	100.000

5. Model Free Evidence

In this chapter, I visualize the data at the weekly aggregated level. There are data visualizations for all six dependent and independent variables that will be analyzed in the following chapters. All figures show the development of the variables from 2018 to 2022, with a vertical dotted line in March 2020 indicating the beginning of the COVID pandemic. The same figures with the timeline starting in 2015 to 2022 can be found on appendix A3.

Figure 3 shows the development of the COVID stringency index over time. The data represents the weekly average of the stringency index for all 31 countries in the dataset. As can be seen, the highest spike of restrictions occurred between March and April of 2020, reaching the highest stringency index level at a value above 80. The average stringency is gradually reduced in the summer of that year, showing the ease of the first wave. Another slight increase in restriction around winter 2020 signaling the second wave of infections. As restrictions start to decrease in 2021, the average index can become smaller over that year, reaching an average value of a little above 40 by the end of that year.

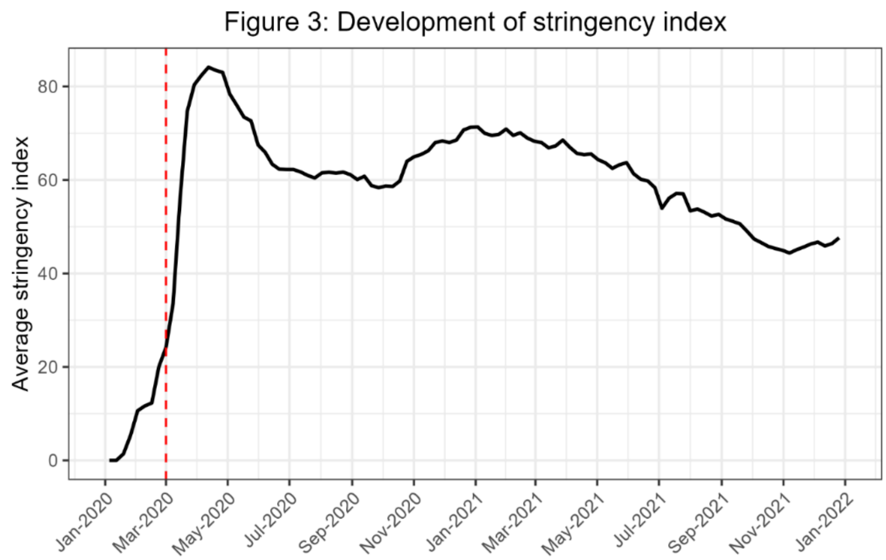


Figure 4 shows the development of the average user rating given to books over time. There is a clear upward trend from 2018 to the end of 2021. This figure doesn't give a strong hint as to whether COVID restrictions influenced the rating users gave, especially since there is an upper trend coming from years prior and there's no clear change at the start of the pandemic.

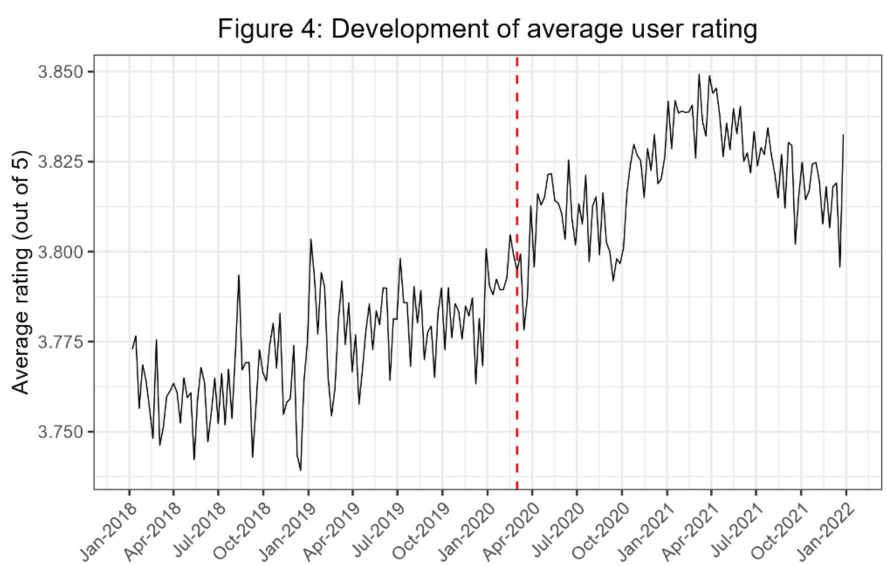


Figure 5 shows the average days a user takes to read a book. There are seasonal dips around Christmas showing that people take shorter to read around this time of the year. There is also quite a visible dip in the month of April of 2020, at the beginning of the COVID restrictions and the higher period of restrictions in the data. Christmas of 2020, another time of high levels of restrictions, also show the largest dip in all the visualization.

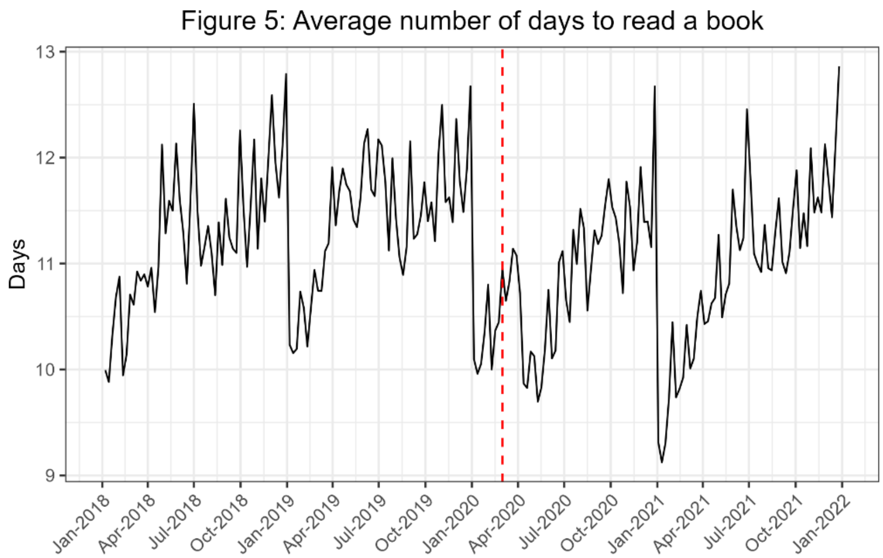


Figure 6 looks at the average age of a book read each week. The two highest peaks in the graph are at the start of the pandemic, around March and April of 2020, and around the holidays of the same year, another time of relatively high restriction rates. The average age of a book read can be seen as slightly higher in 2020 than in the rest of the years in the data.

Figure 6: Development of average age of books read

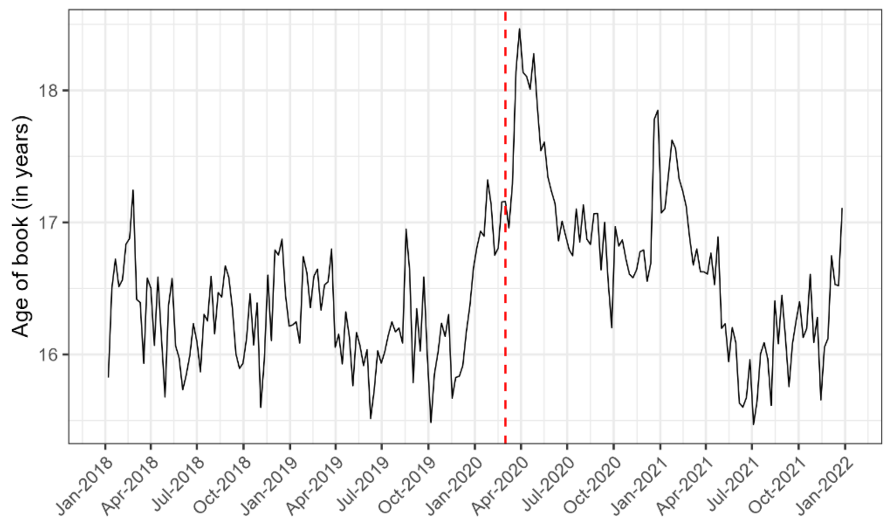
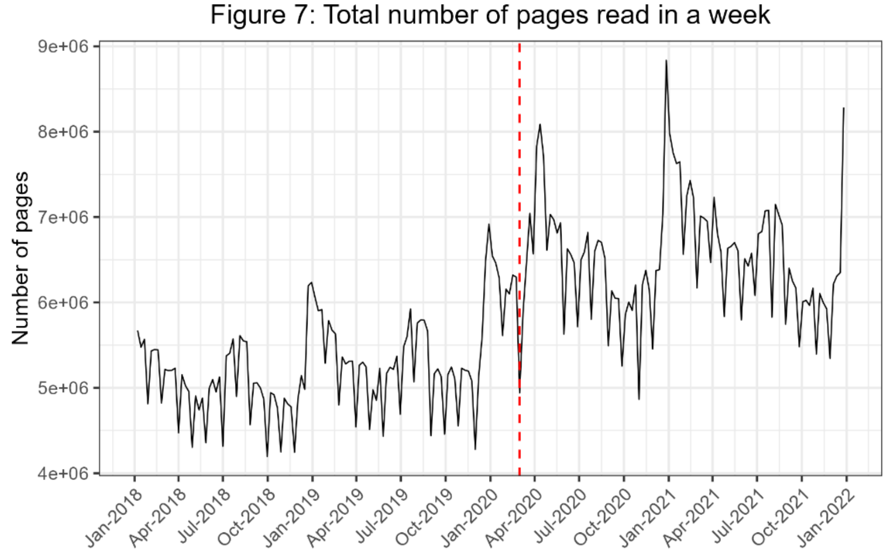


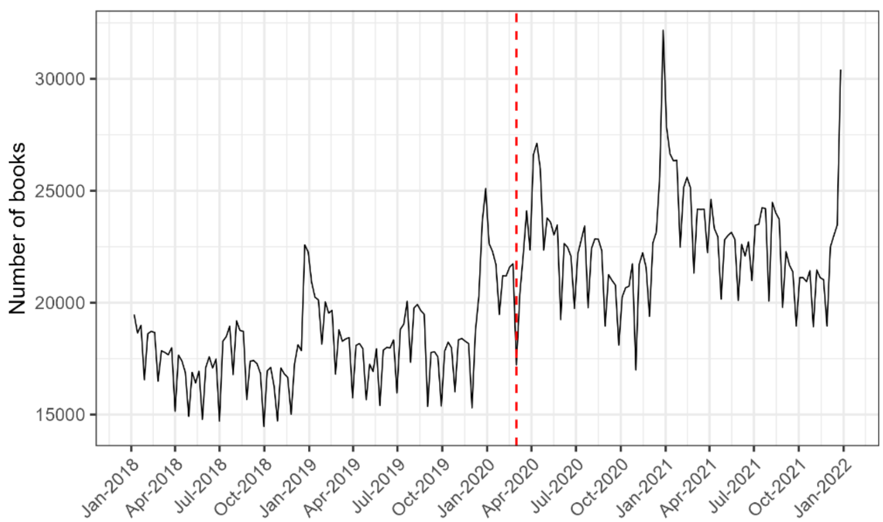
Figure 7 looks at the number of weekly pages for all users in the data. Seasonal consumption patterns around the year-end holidays are visible yearly, along with slight yearly increases in the number of pages read. A spike in April of 2020, at the start of the pandemic, is quite visible, with more pages read than in the holidays of 2019. The highest spike in the graph can also be seen for Christmas of 2020, hinting towards both seasonal patterns and COVID restrictions playing a role in the high increase. Through 2020 and 2021, an increase in almost all months compared to the rest of the years is visible.

Figure 7: Total number of pages read in a week



Like Figure 7, Figure 8 shows the sum of books read by users in each week of the year. As in Figure 5, seasonal consumption patterns are present, and a spike in April 2020 is observable. Like in Figure 5, Figure 6 shows its highest spike in Christmas of 2020, hinting towards a mix of seasonal consumption patterns and COVID restrictions playing a role.

Figure 8: Total number of books read in a week



To summarize, almost all figures depicting dependent variables except for the average user rating seem to have at least a slight effect due to COVID restrictions. The visualizations hint towards users reading slightly faster, older books, more pages, and more books as COVID restrictions increase. To test whether these visualizations show a causal relationship, the data is fitted into the model presented in the following chapters.

6. Regression models

I use a panel regression in the form of an OLS Fixed Effects regression to determine the effects of a shock of an increase in leisure time on book consumption patterns. I present a model where I use the covid categorization of light restrictions, closing of some of society and closing of all but essential parts of society as dependent variables, where each of these variables take a value of 1 if present, and a value of 0 if not.

Appendix A4 includes an alternative model where the independent variable is a continuous variable that ranges from 0 to 100.

6.1 Model showing the differential effect of COVID stringency

Dividing the level of restrictions into four levels instead of directly using the stringency index as a continuous variable from 0 to 100 allows me to correct for the fact that the stringency index is measured on an ordinal scale, which means that the magnitude of the changes in the intervals might not be equal across the index. For example, an increase from 0 to 20 might mean that there is a lot of advertisement to prevent COVID, but not actual measures like closing restaurants, which might be represented on an increase from 40 to 60. This means that an increase from 0 to 20 would not impact society as much as an increase from 40 to 60. Dividing the stringency index into intervals representing different restriction levels solves this problem.

I adapted the stringency index to have four categories of lockdowns. The four categories are: no restrictions, light restrictions, closing of some of society, and closing of all but essential of society. An explanation to how I arrive to these four categories can be found in chapter 4.3. The model used is the following:

$$log\, Y_{ijcpy} = \beta_0 + \sum_{k=1}^{4} \beta_{1(k)}\, Adapted\ Stringency\ Index_{(k)_{y,p,c}} + \lambda_p + \gamma_y + \alpha_i + \mu_j + \varepsilon_{ijcpy}$$

Where K represents each of the four restriction levels presented and the subscripts represent a given user "i," book "j," country "c," biweek "p", and year "y." Each of the categories of restrictions, light restrictions, closing of some of society and closing of all of society take a value of 1 if that level of restriction is present, and a value of 0 if there are no restrictions. For example, when evaluating the light level of restrictions, that variable will take a value of 1 if present and a value of 0 if there are no restrictions. This allows to evaluate how book consumption patterns might change at different levels of restrictions. This model evaluates the dependent variables "user rating" and "reading time per book".

The explanation and rationale for each fixed effect are the following:

- Biweek Fixed Effect: It is constructed with a variable with values ranging from 0 to 26, representing each a group of two weeks of a given year. The main goal of this fixed effect is to account for seasonality. An example for seasonality patterns is that more users read books closer to Christmas or during the summer than in other months of the year. Table 1 shows that users take an average of 11.18 days to read a book, with a standard deviation of 23.68 days. Therefore, from Table 1, I can conclude that the measure of a biweek is representative to allow for users to have enough time to read a book in that period.
- Year Fixed Effect: It is constructed with a variable representing each year in the data. The purpose of this fixed effect is to correct for yearly changes in the data. For instance, the number of books users read on the platform has increased yearly. This variable captures average yearly changes in the data across all users that are not attributed to the stringency restrictions.
- User FE: Users may have different characteristics affecting how they consume books. The characteristics controlled for by this fixed effect are the following: country, number of ratings given, the average rating given to books, number of reviews posted, number of books read, number of friends, age, gender, date joined, average books read per day, and total books left to read in their library. At the same time, Zhang & Godes (2018) describe how users with more friends on the platform give higher ratings to books. Also, Thelwall & Kousha (2017) found that older users are more active than newer users. This shows that specific user characteristics are correlated with some of the dependent variables in the model. This fixed effect looks to eliminate any individual characteristics that might bias the estimate of the interest variable.

- Book FE: Books may have different characteristics, making users consume these differently. The characteristics controlled for by this fixed effect are the following: average rating of the book, the number of pages it has, the number of ratings, the year it was published, its age, and the book's name and author. This fixed effect looks to eliminate any book characteristics that might bias the estimate of the interest variable.

Clustering standard errors allows to account for potential correlation and control for unobserved country-specific factors that may be correlated with the regressors and the outcome variables. This correlation belongs to the error term of the regression model. Failure to account for clustering may bias standard errors and p-values, leading to incorrect inferences and conclusions (Wooldrige, (2018)). It was decided to use clustered standard errors at the country and biweek level. The error variance within a country and a two-week period may be correlated. Clustering at the country level is appropriate, as COVID restrictions simultaneously affect everyone in the country. Clustering at the biweek level is appropriate because the error term may be correlated within the same biweek in different years due to unobserved factors that affect the outcome variable similarly each year, such as weather patterns, holidays, or other seasonal effects. Clustering standard errors at the biweek level can account for this potential correlation of the error variance within each biweek across different years.

6.2.1 Alterations to the model: absence of book FE

The level of aggregation for the remaining three dependent variables, which are "age of book," "Number of pages read," and "Number of books read," changes from the user-book level to the user level. Therefore, the remaining dependent variables do not vary at the book level but only at the individual and time level. As a result, I need to amend the model to not include the book fixed effects.

$$log\, Y_{ijcpy} = \beta_0 + \beta_1 \sum_{k=1}^{4} \beta_{1(k)}\, Adapted\ Stringency\ Index_{(k)_{y,p,c}} + \lambda_p + \gamma_y + \alpha_i + \varepsilon_{ijcpy}$$

The subscripts and fixed effects in this model have the same meaning as the original model of the previous section. Like the original model, the model described in this section also clusters standard errors at the country and biweek level.

7. Results

This section presents the results of fitting the data in the models of chapter 6. The results of the model that shows the differential effect of COVID stringencies are presented at both the book and user levels. The three levels of restrictions that will be shown are light restrictions, closing of some of society, and closing of all but essential of society. The results to the alternative model with a continuous independent variable are presented in appendix A5. As in marketing and economics research, the results are interpreted as significant at the 5% level.

Table 2 shows the results for the model showing the differential effects of COVID restrictions at the individual-book level. I start evaluating the read time per book mode. The results show that when restrictions are light, this does not significantly affect users' reading speed. As restrictions increase to the level of closing some of society, users read significantly faster, at about 6.7% faster than if there were no restrictions. When society shuts down outside what is essential, people read 14.5% significantly faster than without restrictions. To put the results into perspective, when some of society closes, users take 0.75 days less to complete a book, and when all of society closes, users take 1.57 days less to complete a book, compared to the average of 11.18 days. After using the linear hypothesis test from the car package in R, both coefficients, 6.7% and 14.5%, are significantly different from each other. It can be debated whether the book fixed effects are necessary for this variable; because of this, it was decided to present results for both book FE and no book FE. All in all, book FE matter for the user's reading speed. Some books take longer to read than others, and this is because of the length, genre, or way the book is written. Because of this, book FE should be included for this variable.

Turning to the effect of COVID restriction on the average age of a book read. When restrictions are light, there are no significant effects on this variable. As restrictions increase and reach a level of closing some of society, people read books that, on average, are about 2.7% older. When all but essential parts of society are closed, people read books 5.7% older than when there are no restrictions. To put the results into perspective, when some of society closes, users read books that are 0.43 years older, and when all of society closes, users read books that are 0.91 years older, compared to the average of 16.04 years. Both coefficients, 2.7% and 5.7%, are significantly different from one-another.

On the last variable of the table, we look at the effects of COVID restrictions on the average rating given by users to books. Light and strong restrictions seem not to influence the way users rate books. There is a slight significant increase of about 0.3% in the user rating of books when some of society is closed. Even though this is significant, the increase of 0.3% on a scale of 5 stars is small, at about a 0.015 increase in stars, meaning that the increase is negligible. Hence, restrictions seem not to influence the way users rate books.

Table 2: Differential effects of COVID stringency at the book level

	Log (Reading time per book)	Log (Reading time per book)	Log (Age of book)	Log (User rating)
Light restrictions	-0.001	-0.010	-0.003	0.001
	(0.013)	(0.011)	(0.011)	(0.001)
Closing of some of society	-0.067***	-0.052***	0.027*	0.003**
	(0.015)	(0.011)	(0.010)	(0.001)
Closing of all but essential of society	-0.145***	-0.103***	0.057***	0.000
	(0.030)	(0.022)	(0.012)	(0.001)
Num.Obs.	4235109	4235109	5064903	6046614
R2	0.564	0.242	0.232	0.497
FE: Biweek	X	X	X	X
FE: Year	X	X	X	X
FE: User	X	X	X	X
FE: Book	X			X

Note: Clustered standard errors at the biweek and country level. (+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$)

Table 3 displays the results of the model using differentiated levels of restrictions at the user level. The findings reveal that light restrictions do not affect the average number of pages read. When some of society is required to close, there is a 6.9% increase in the average number of pages read by users. When all but essential parts of society are closed, on average, users significantly read more at an increase of about 11.6%. To put the results into perspective, when some of society closes, users read about 43 more pages in a span of two weeks, and when all of society closes, users read about 72 more pages in a span of two weeks, compared to the average of 619 pages in the same span of time. Both coefficients, 6.9% and 11.6%, are significantly different from one-another.

Similarly, the number of books read by users is not significantly affected by light restrictions. However, when some of society closes, there is a statistically significant 6.5% increase in the number of books read. When all but essential parts of society are closed, users read significantly more books, with a 10.4% increase compared to no restrictions. To put the results into perspective, when some of society closes, users read about 0.14 more of a book in two weeks, and when all of society closes, users read about 0.23 more of a book , compared to the average of 2.163 books read in two weeks. Both coefficients, 6.5% and 10.4%, are significantly different from one-another.

Table 3: Differential effect of COVID stringency at the user level

	Log (Number of pages read by a user)	Log (Number of books read by a user)
Light restrictions	0.048+	0.049+
	(0.027)	(0.027)
Closing of some of society	0.069*	0.065*
	(0.025)	(0.025)
Closing of all but essential of society	0.116**	0.104**
	(0.034)	(0.029)
Num.Obs.	2514974	2718998
R2	0.286	0.357
FE: Biweek	X	X
FE: Year	X	X
FE: User	X	X
FE: Book		

Note: Clustered standard errors at the biweek and country level. (+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$)

With these results it is possible to look back at the hypotheses proposed in chapter 3.4. The results support the first hypotheses, *book consumption increases.* Three models support this hypothesis: the models for number of pages and books read by a user, and the model that evaluate the reading time per book These models show that the higher restrictions level is, the higher the consumption of books. Users read significantly more pages, books, and take less days to read these books as leisure time increases. Hypothesis two, *users will read older books,* is also supported. As restrictions get tighter, users read significantly older books. Finally, hypotheses three, *users will not change the way they rate books,* is also supported. At almost all levels of restrictions, these do not have an effect on the way people rate books, and if they do, it is a negligible effect.

8. Discussion and conclusion

This thesis studied the effects of shocks to leisure time in the form of COVID restrictions on book consumption. I looked to answer the research question "*How do consumers leverage extra leisure time to book-related consumption on Goodreads?"* and the sub questions *"Do consumers of the platform Goodreads predominantly adopt an exploration or preference discovery mechanism when there is a positive shock to their leisure time?"* and "*Does extra leisure time make consumers more critical of the products they consume?".* I stated three hypotheses predicting that, *book consumption will increase, users will read older books, and users will not change the way they rate books as their leisure time increases.* I was able to answer these questions and hypotheses by using an OLS Fixed Effects regression that fitted data from the platform Goodreads and the Oxford COVID-19 Government Response Tracker.

I found the effects of four different levels of restrictions on five different dependent variables that study book consumption patterns. To answer the main research question, consumers leverage extra leisure time by consuming more pages and books at a faster rate. Regarding the first sub-question, while it is not possible to confirm nor deny the presence of an exploration mechanism, there is evidence of a preference discovery mechanism as users consume more books that are also older. On the second sub-question, extra leisure time does not affect the way people perceive the products they consume, as extra leisure time does not influence the way users rate books. The three hypotheses I proposed are all supported. First, book consumption increases in the form of more pages and books read, and users also read these books faster. Second, as leisure time increases, users read older books. Third, extra leisure time does not affect the way people rate books. All effects found are only significant at the levels of restrictions that force some or all of society to close, with stronger restrictions intensifying all results.

My research brings implications to the field of economics and for managers. For economics, it contributes to understanding consumer behavior during periods of shocks or disruptions to society. Understanding how consumers leverage their book consumption patterns can contribute to the broader field of consumer behavior research that provides insights for economists when studying the effects of shocks to consumption patterns. For managers, my research brings implications for marketing strategies for book-related products. Since consumers tend to read more books and older titles during periods of increased leisure time, managers can promote older books to satisfy consumer preference during this period of time. The findings of my thesis can also reassure managers that extra leisure time does not affect the way users rate books, which is likely to be the case for other products that are not books. This can allow managers to make informed when promoting or developing new products during times of shocks to leisure time. Lastly, my research provides tools to managers of the book industry to adapt their business strategies to anticipate changes in consumer behavior at the time of leisure time shocks. For example, managers can adapt their supply of books, marketing, and distribution strategies to meet the changes in consumer behavior during times of shocks to people's leisure time.

My thesis presents some limitations that opens different avenues for future research. My thesis is not able to disentangle the mechanisms behind the increase in book consumption. There are multiple mechanisms that could be driving the increase in book consumption. While there is evidence supporting the preference discovery mechanism, there may be others in the background, such as an exploration discovery mechanism, where users seek new products, such as watching more TV shows, or income expansion path, where an increase in income and leisure time could increase book consumption. Future research could look to construct econometric models that could disentangle these mechanisms. Another limitation and potential research avenue is that my research focuses on books as a whole and does not differentiate the changes in the genres of each book. It may be the case that the consumption of fiction books increases significantly more compared to non-fiction books. Future research could delve into more detail on the effects of the changes in books consumption in the different types of books. Finally, my research shows that book consumption increases as leisure time increases, but it is not clear as to whether this increase lingers once the restrictions disappear. It might be the case that book consumption is significantly higher after the exposure to leisure time shocks. Future research could study the long-term effect to the exposure to these leisure time shocks.

9. References

Alisa Flood. (2013, April 12). *Amazon purchase of Goodreads stuns book industry*. https://www.theguardian.com/books/2013/apr/02/amazon-purchase-goodreads-stuns-book-industry

Chen, H., Qian, W., Wen, Q., Agarwal, S., Davis, S., Fang, H., He, Z., Hsieh, C.-T., Huang, Y., Park, A., Qin, Y., Song, Z., Sulaeman, J., Tang, D., Wang, J., Wang, P., Yang, L., Yeung, B., Zaldokas, A., … Zhu, H. (2020). *The Impact of the COVID-19 Pandemic on Consumption: Learning from High Frequency Transaction Data * We benefit from valuable comments by The Impact of the COVID-19 Pandemic on Consumption: Learning from High Frequency Transaction Data*.

Chetty John Friedman Nathaniel Hendren Michael Stepner, R. N., Autor, D., Chodorow-Reich, G., Farhi, E., Furman, J., Hamilton, S., Hurst, E., Jaravel, X., Katz, L., Lange, F., Saez, E., Straub, L., Yagan, D., Rajagopal, A., Anton Libsch, L., Taska, B., Arun Natesan, E., Palaniappan, R., Ray Sandza, H., … Sharma, S. (2020). *Stabilization Policies Affect Spending and Employment? A New Real-Time Economic Tracker Based on Private Sector Data." We thank*. http://www.nber.org/papers/w27431

Curcic, D. (2023, January 26). *Global Book Sale Statistics*. https://wordsrated.com/global-book-sales-statistics/

Ding, A. W., & Li, S. (2019). Herding in the consumption and purchase of digital goods and moderators of the herding bias. *Journal of the Academy of Marketing Science*, *47*(3), 460–478. https://doi.org/10.1007/s11747-018-0619-0

Emma Charlton. (2020, April 30). *Coronavirus escapism: book sales surge during lockdown*. https://www.weforum.org/agenda/2020/04/coronavirus-escapism-book-sales-surge-covid-19/

Federation of European Publishers. (2021). *Consequences of the COVID-19 crisis on the book market-An overview of 2020*.

Goodreads. (2023). *About Goodreads*. https://www.goodreads.com/about/us

Guren, C., McIlroy, T., & Sieck, S. (2021). COVID-19 and Book Publishing: Impacts and Insights for 2021. *Publishing Research Quarterly*, *37*(1). https://doi.org/10.1007/s12109-021-09791-z

Hale, T., Angrist, N., Goldszmidt, R., Kira, B., Petherick, A., Phillips, T., Webster, S., Cameron-Blake, E., Hallas, L., Majumdar, S., & Tatlow, H. (2021). A global panel database of pandemic policies (Oxford COVID-19 Government Response Tracker). *Nature Human Behaviour*, *5*(4), 529–538. https://doi.org/10.1038/s41562-021-01079-8

Jeffrey M. Wooldrige. (2018). *Introductory Econometrics: a modern approach* (Cengage, Ed.; 7th ed., pp. 426–494).

Jonty Bloom. (2021, October 7). *Booksellers hope soaring sales will continue as we read more*. https://www.bbc.com/news/business-58802805

Lee, I., & Tipoe, E. (2021). Changes in the quantity and quality of time use during the COVID-19 lockdowns in the UK: Who is the most affected? *PLoS ONE*, *16*(11 November). https://doi.org/10.1371/journal.pone.0258917

Liao, H. L., & Liu, S. H. (2022). Integrating Information Technology and Marketing to increase e-Book consumption. *Electronic Commerce Research*. https://doi.org/10.1007/s10660-022-09585-1

Michelle. (2022, October 28). *Goodreads: A Platform for Readers and Authors*. https://d3.harvard.edu/platform-digit/submission/goodreads-a-platform-for-readers-and-authors/

Oblander, S., & Mccarthy, D. M. (n.d.). *Estimating the Long-Term Impact of Major Events on Consumption Patterns: Evidence from COVID-19*. https://ssrn.com/abstract=3836262

Richard Stuart-Turner. (2021, January 28). *Printed book sales pass 200m in 2020 despite pandemic*. https://www.printweek.com/news/article/printed-book-sales-pass-200m-in-2020-despite-pandemic

Sim, J., Cho, D., Hwang, Y., & Telang, R. (2022). Frontiers: Virus Shook the Streaming Star: Estimating the COVID-19 Impact on Music Consumption. *Marketing Science*, *41*(1), 19–32. https://doi.org/10.1287/mksc.2021.1321

Statista. (n.d.). *Number of registered members of Goodreads*. 2023. Retrieved 19 June 2023, from https://www.statista.com/statistics/252986/number-of-registered-members-on-goodreadscom/

Thelwall, M., & Kousha, K. (2017). Goodreads: A social network site for book readers. *Journal of the Association for Information Science and Technology*, *68*(4), 972–983. https://doi.org/10.1002/asi.23733

Waldfogel, J., & Kappel Chair, F. R. (2023). *THE WELFARE EFFECT OF GENDER-INCLUSIVE INTELLECTUAL PROPERTY CREATION: EVIDENCE FROM BOOKS*. http://www.nber.org/papers/w30987

Zhang, Y., & Godes, D. (2018). Learning from online social ties. *Marketing Science*, *37*(3), 425–444. https://doi.org/10.1287/mksc.2017.1076

10. Appendix

A1: Countries in the data

The list of 31 countries found in the dataset are the following: Egypt, Indonesia, Italy, Pakistan, Argentina, Portugal, Lithuania, Australia, Poland, Netherlands, United Arab Emirates, Finland, Norway, United Kingdom, Russia, Malaysia, Romania, India, Philippines, Mexico, Bangladesh, Swede, Nepal, Turkey, Singapore, Estonia, Bulgaria, Latvia, Austria, Panama, Venezuela.

A2: Construction of the adapted stringency index

The categories and indexes used to construct the stringency index, with their respective value range (in parenthesis), are the following:

- C1 (0-3): Closing of schools and universities
- C2 (0-3): Closing of workplaces
- C3 (0-2): Cancelling public events
- C4 (0-4): Limits on gatherings
- C5 (0-2): Closing of public transportation
- C6 (0-3): Orders to "shelter-in-place" and otherwise stay at home
- C7 (0-4): Restrictions on international movement between cities/regions
- C8 (0-4): Restrictions on international travel
- H1 (0-2): Presence of public information campaigns

Flags are used to determine the geographical impact of a policy. I didn't take any flags into account when constructing the intervals for the stringency index used in this thesis.

The highest value of the indexes for each of the nine ratios are the following:

<u>No restrictions:</u> C1 = 0/3. C2 = 0/3, C3 = 0/2. C4 = 0/4, C5 = 0/2, C6 = 0/3, C7 = 0/2, C8 = 0/4, H1 = 0/2.

Light or recommended restrictions: C1 = 1/3. C2 = 1/3, C3 = 1/2. C4 = 1/4, C5 = 1/2, C6 = 1/3, C7 = 1/2, C8 = 2/4, H1 = 1/2.

Required closing of some of society: C1 = 2/3. C2 = 2/3, C3 = 2/2. C4 = 4/4, C5 = 1/2, C6 = 2/3, C7 = 2/2, C8 = 4/4, H1 = 2/2.

Required closing of some of society: C1 = 3/3. C2 = 3/3, C3 = 2/2. C4 = 4/4, C5 = 2/2, C6 = 3/3, C7 = 2/2, C8 = 4/4, H1 = 2/2.

A3 Visualizations presented in Chapter 5 but with longer time paths

Figure 2 A: Development of average user rating

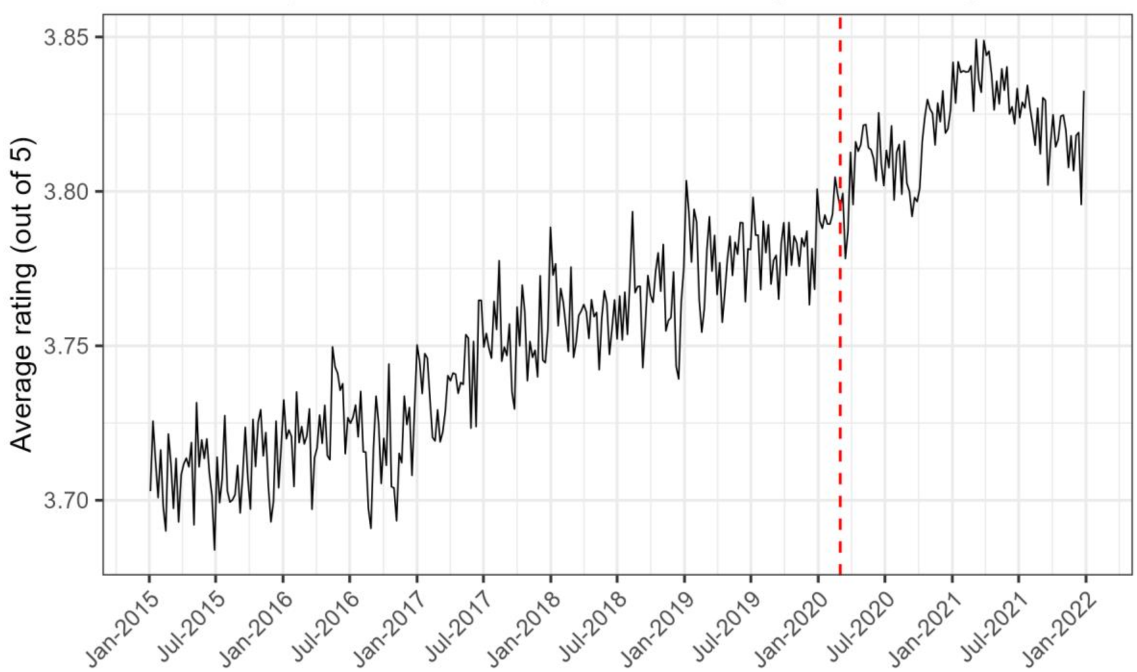


Figure 3 A:Average number of days to read a book

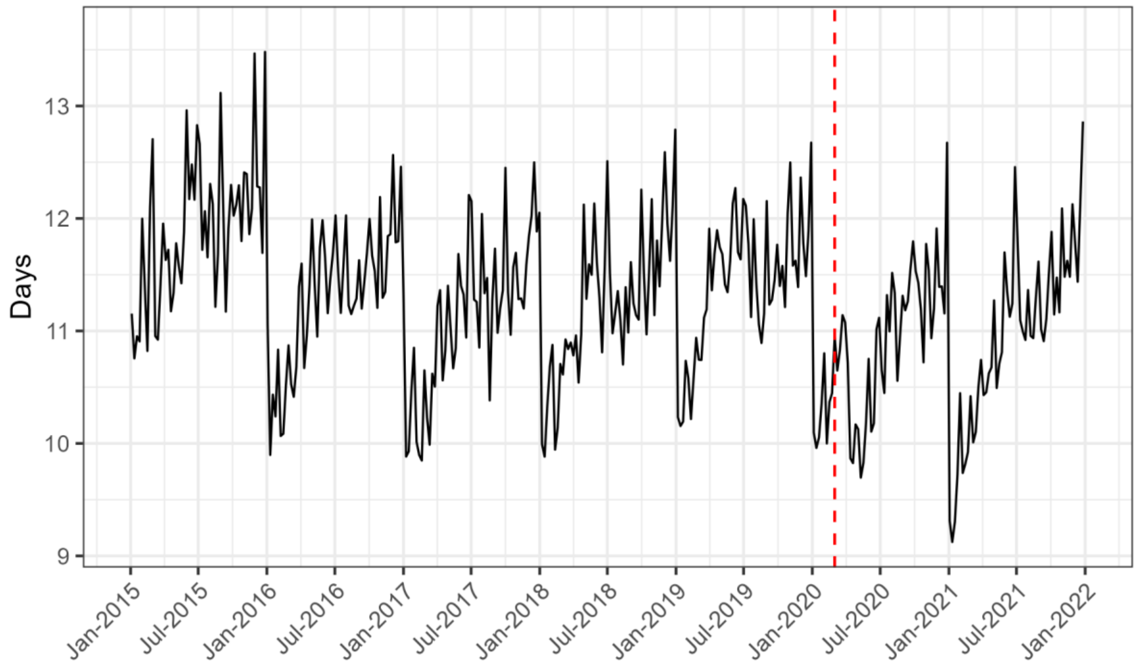


Figure 4 A: Development of average age of books read

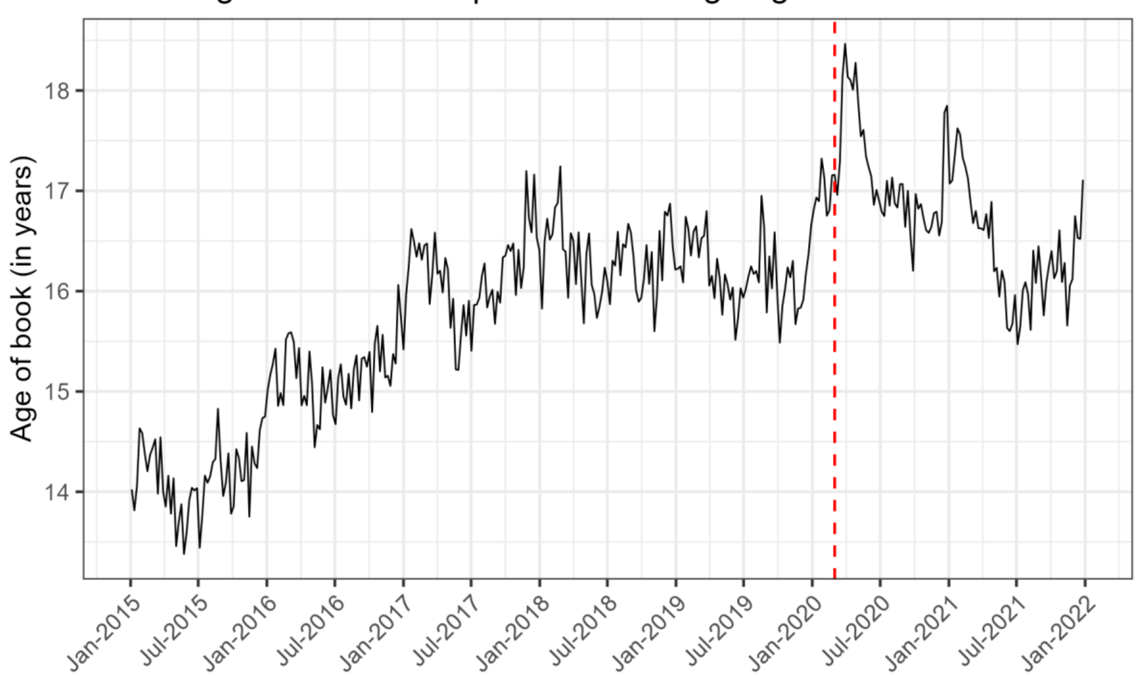


Figure 5 A: Total number of pages read in a week

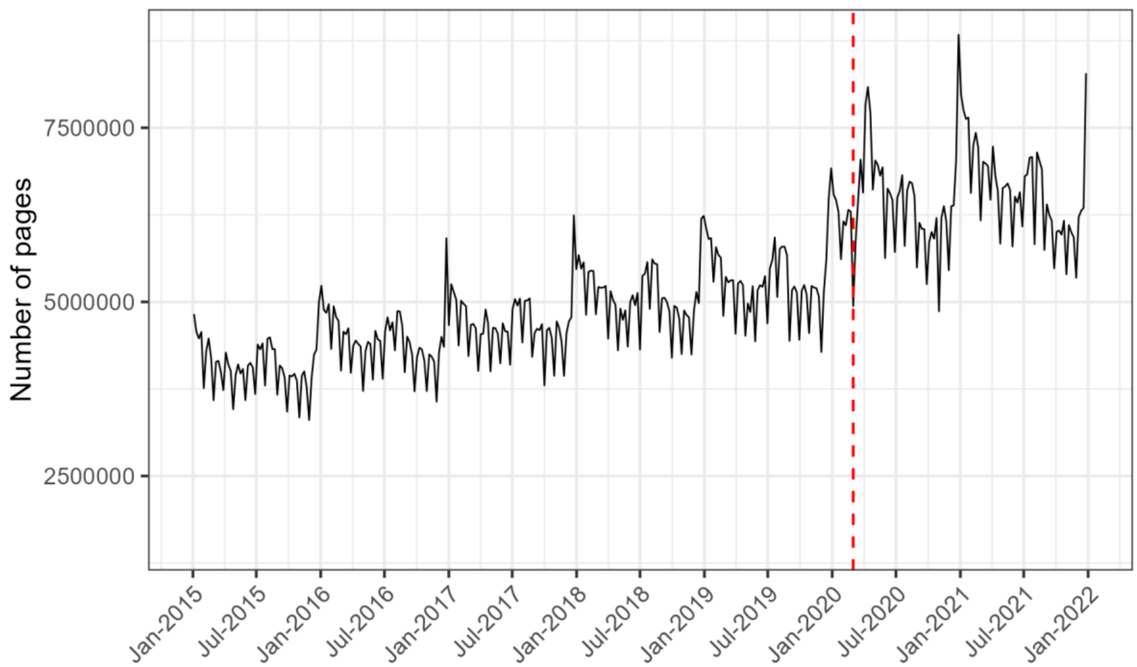
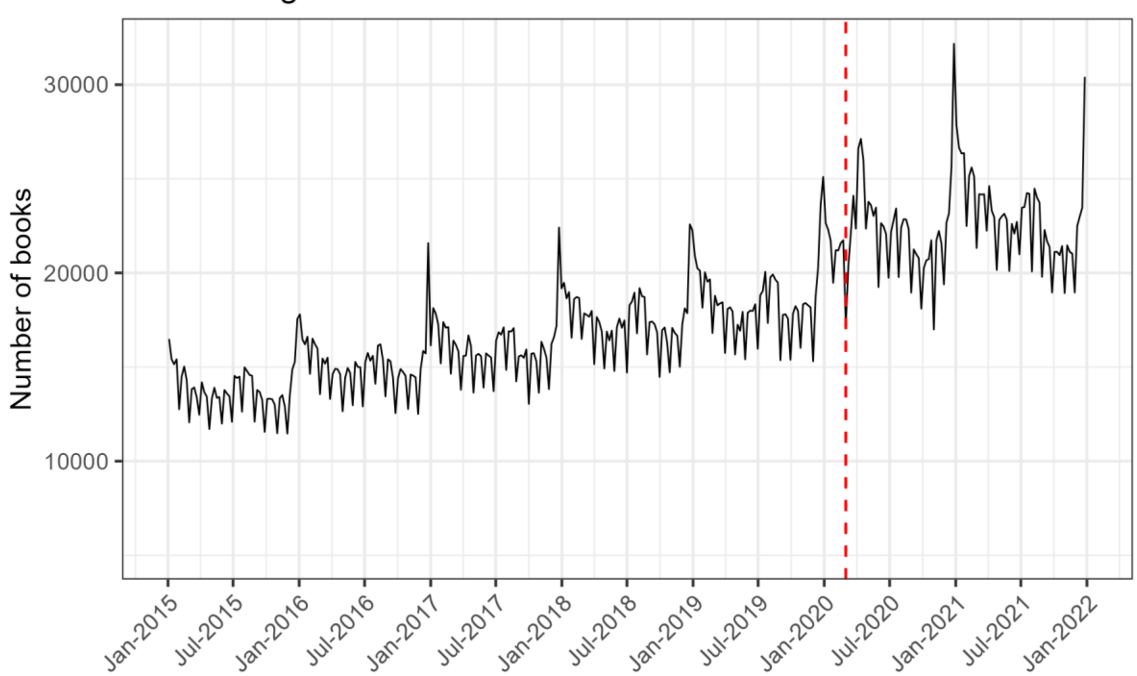


Figure 6 A: Total number of books read in a week



A4: Model with a continuous independent variable

This model counts on the independent variable $Stringency\ Index$, which is continuous ranging from 0 to 100. It is a model with a linear independent variable because of the continuous nature of that variable. The model is the following:

$$log\,Y_{ijcpy} = \beta_0 + \beta_1\,Stringency\ Index_{y,p,c} + \lambda_p + \gamma_y + \alpha_i + \mu_j + \varepsilon_{ijcpy}$$

Where the subscripts represent a given user "i," book "j," country "c," biweek "p", and year "y." The coefficient of interest is divided by 10 for easier interpretation of the coefficients as otherwise too the values would be too small to evaluate. This model evaluates two dependent variables, “user rating” and “reading time per book”. All Fixed Effects are the same as the fixed effects presented in the model of chapter 6. Like that model, this model also uses clustered standard errors at the biweek and country level.

Alterations to the model: absence of book FE

The level of aggregation for the remaining three dependent variables, which are "age of book," "Number of pages read," and "Number of books read," changes from the user-book level to the user level. Therefore, the remaining dependent variables do not vary at the book level but only at the individual and time level. As a result, I need to amend the model to not include the book fixed effects.

$$log\,Y_{ijcpy} = \beta_0 + \beta_1 Stringency\ Index_{y,p,c} + \lambda_p + \gamma_y + \alpha_i + \varepsilon_{ijcpy}$$

The subscripts and fixed effects in this model have the same meaning as the original model of the previous section. Like the original model, the model described in this section also clusters standard errors at the country and biweek level.

A5: Results for the model with a linear independent variable

Table 4 shows the book-level results of the model with a continuous independent variable. Since the stringency index variable is a continuous variable from 0 to 100, I use the 75th percentile of the stringency index, found in table 1, divided by 10 to interpret the results, which is 7.32.

First, the interpretation of the read time per book model. If restrictions increase to the 75th percentile, users take 12.44% less time to read. It can be debated whether the book fixed effects are necessary for this variable; because of this, it was decided to present results for both book FE and no book FE. All in all, book FE matter for the user's reading speed. Some books take longer to read than others, and this is because of the length, genre, or way the book is written. Because of this, book FE should be included for this variable. This is also visible in the results. Turning to the effect of COVID stringency on the average age of a book read, when restrictions reach the 75th percentile level, users read significantly older books, with books read being 5.86% older. Lastly, the effect of COVID stringency on the average rating given by users. On average, COVID restrictions do not significantly affect the ratings users give to books. Even though the effect may be significant, the effect of COVID restrictions on user rating, the coefficient is small at 0.000 that even at the 75th percentile this would not have an effect on a 5-star rating scale.

Table 4: Model with a linear independent variable, book level results

	Log (Reading time per book)	Log (Reading time per book)	Log (Age of book)	Log (User rating)
COVID stringency	-0.017***	-0.011***	0.008***	0.000*
	(0.003)	(0.002)	(0.001)	(0.000)
Num.Obs.	4235109	4235109	5064903	6046614
R2	0.564	0.242	0.232	0.497
FE: Biweek	X	X	X	X
FE: Year	X	X	X	X
FE: User	X	X	X	X
FE: Book	X			X

Note: Clustered standard errors at the biweek and country level. (+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$)

Table 5 shows the results for the user-level variables. We turn to the effect of COVID restrictions on the average number of pages read by a user. At the 75th percentile level of restrictions, on average, people read significantly more pages, at about 6.59%, than when there are no restrictions. On the effect of COVID restrictions on the average number of books read, at the 75th percentile level of restrictions, on average, people read significantly more books, at about 5.12%, than when there are no restrictions.

Table 5: Model with a linear independent variable, user level results

	Log (Number of pages read by a user)	Log (Number of books read by a user)
COVID stringency index	0.009***	0.007***
	(0.002)	(0.002)
Num.Obs.	2514974	2718998
R2	0.286	0.357
FE: Biweek	X	X
FE: Year	X	X
FE: User	X	X
FE: Book		

Note: Clustered standard errors at the biweek and country level. (+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$)